

O'REILLY®

Applied Embedded Electronics

Design Essentials
for Robust Systems

Early
Release

RAW &
UNEDITED



Jerry Twomey

VISIT...

LANZAROTE
Caliente.COM

Applied Embedded Electronics

Design Essentials for Robust Systems

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

Jerry Twomey

Applied Embedded Electronics

by Jerry Twomey

Copyright © 2023 Gerald Twomey. All rights reserved.

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North,
Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (<http://oreilly.com>). For more information, contact our corporate/institutional sales department: 800-998-9938 or corporate@oreilly.com.

Acquisitions Editor: Brian Guerin

Development Editor: Angela Rufino

Production Editor: Christopher Faucher

Interior Designer: David Futato

Cover Designer: Karen Montgomery

- December 2023: First Edition

Revision History for the Early Release

- 2022-12-08: First Release
- 2023-01-25: Second Release
- 2023-03-02: Third Release
- 2023-05-10: Fourth Release
- 2023-06-23: Fifth Release

See <http://oreilly.com/catalog/errata.csp?isbn=9781098144791> for release details.

The O'Reilly logo is a registered trademark of O'Reilly Media, Inc. *Applied Embedded Electronics*, the cover image, and related trade dress are trademarks of O'Reilly Media, Inc.

The views expressed in this work are those of the author, and do not represent the publisher's views. While the publisher and the author have used good faith efforts to ensure that the information and instructions contained in this work are accurate, the publisher and the author disclaim all responsibility for errors or omissions, including without limitation responsibility for damages resulting from the use of or reliance on this work. Use of the information and instructions contained in this work is at your own risk. If any code samples or other technology this work contains or describes is subject to open source licenses or the intellectual property rights of others, it is your responsibility to ensure that your use thereof complies with such licenses and/or rights.

978-1-098-14473-9

Chapter 1. Essential Concepts

A NOTE FOR EARLY RELEASE READERS

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

This will be the 1st chapter of the final book.

If you have comments about how we might improve the content and/or examples in this book, or if you notice missing material within this chapter, please reach out to the editor at arufino@oreilly.com.

An Introduction to electronics gives the basics of Ohms law, Kirchhoff’s current & voltage laws, current and voltage sources, how resistors, capacitors and inductors function, and similar. A typical curriculum will treat these things as ideal to clearly convey the basics. Unfortunately, academic simplifications can be misleading. Real world applications frequently run into problems when ideal concepts fall short. Nothing is ideal.

This chapter takes a brief look at some of the “less than ideal” components seen in real world applications, how they differ from ideal devices, and where designers need to pay attention to the limitations of non-ideal devices. This chapter is about building an awareness of problems that are commonly encountered; while the rest of the book develops solutions to, or, strategies to avoid, these issues. In order to solve any problem, the designer needs to be aware of the problem first. Again, nothing is ideal.

Remember That: Basic Electronics

There are many textbooks that cover fundamental electronic design from scientific, engineering and hobbyist perspectives. The intention here is not basic electronics, but going into the important things needed to design a reliable electronic system. The following illustrations serve as a quick benchmark. If you can look through these two panels of topics, and be reasonably familiar with them you should have the needed background to work with the material presented here (**Figure 1-1** and 1-2).

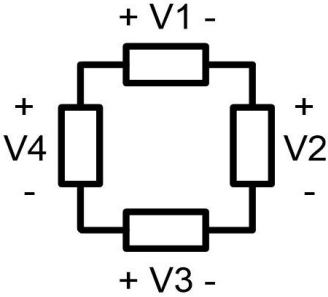
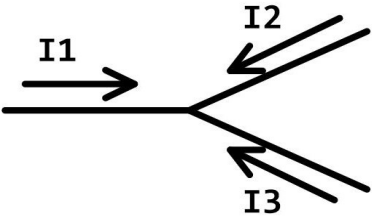
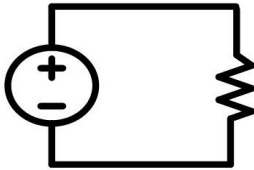
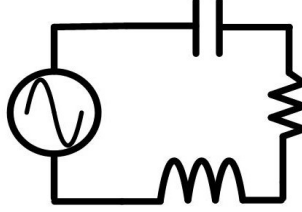
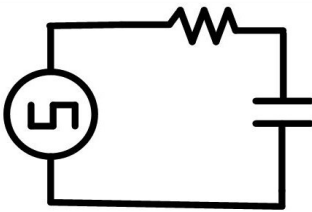
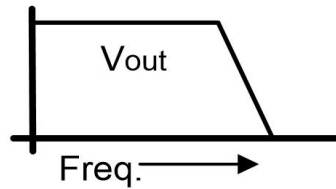
Current (Amperes)	Voltage (Volts)	Charge (Coulombs)	Resistance (Ohms)
Ohm's Law (DC): $V = IR$ Ohm's Law (AC): $V = IZ$			
Power (Watts): $P = VI = I^2 R = V^2 / R$			
Capacitors: $I = C \, dV/dT$ $Q = CV$ $ Z_C = 1 / (2 \pi f C)$			
Inductors: $V = L \, dI/dt$ $ Z_L = 2 \pi f L$			
Kirchhoff's Voltage Law: $\sum V = 0$			
Kirchhoff's Current Law: $\sum I = 0$			
Steady State DC Systems:			
Steady State AC Systems:			
Transient Signals:			

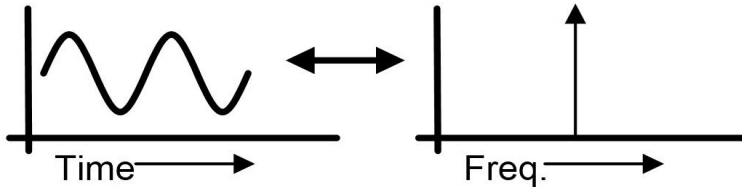
Figure 1-1. Basic concepts (1)

Decibel: $\text{dB} = 20 \log_{10} (V_1/V_2) = 10 \log_{10} (P_1/P_2)$

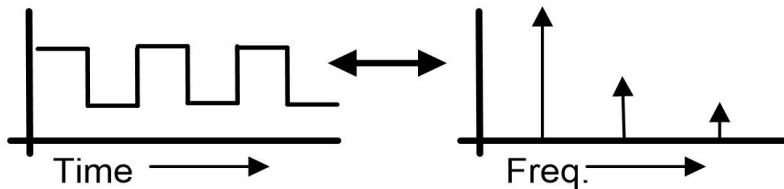
Bode Plots &
Frequency
Response



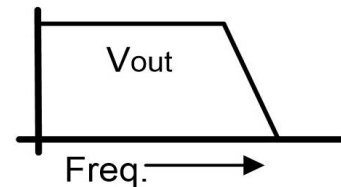
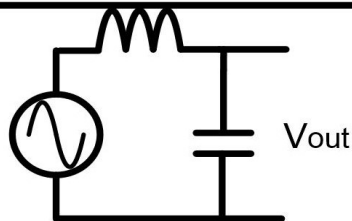
Time Domain vs.
Frequency Domain,
Fourier Equivalents



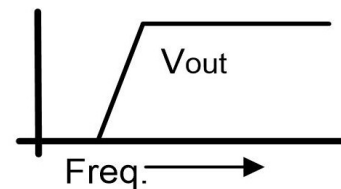
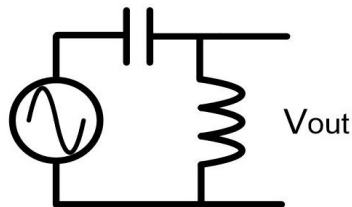
Harmonics in
Frequency Domain



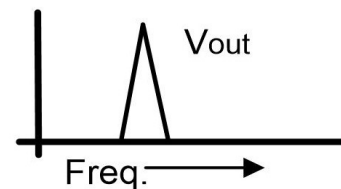
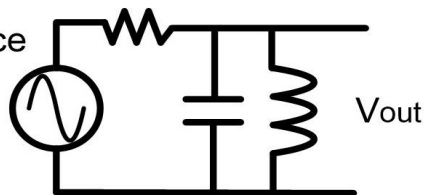
LC Resonance
Low Pass
Filter:



LC Resonance
High Pass
Filter:



LC Resonance
Band Pass
Filter:



RC Time
Constant:

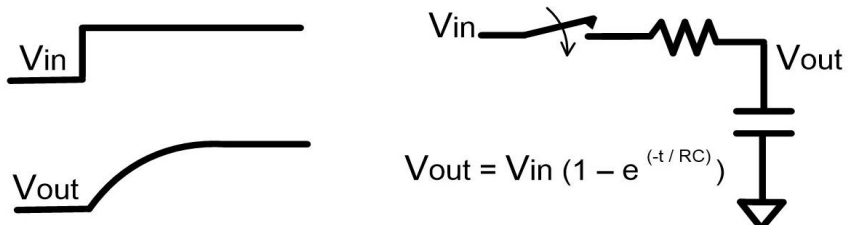


Figure 1-2. Basic concepts (2)

If some items are unfamiliar, reference books, online search, or many online tutorials exist. Many sources cover the basics, but that is not the focus here.

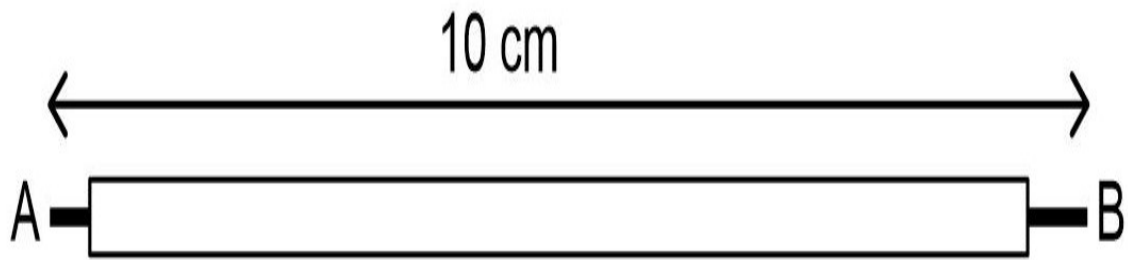
Ideal Simplifications of Academia

Modern electronic systems are predominantly digital, but most design problems are analog in origin. Items such as noise, signal integrity, power stability, electromagnetic interference (EMI), and connection impedance are common problems. These issues can quickly degrade, or render nonfunctional, any electronic system. Digital systems are more tolerant of these issues, but fully digital devices can be broken by analog limitations.

When explaining electronics, a simplified representation gives a quick idea of how something should work. That “first order” model is useful for simplicity and illustration, but frequently leaves out important details. That simple model is often an incomplete model and many of the “second order effects” left out can affect device performance significantly. Becoming aware of a more detailed model helps lay a foundation for better design. The techniques presented here all pay attention to second order effects, and provide methods to deal with them.

Interconnections

As a start, the idea of connecting things together needs a closer look (**Figure 1-3**):



- Solid Copper Wire
- 24 AWG
- Placed 5mm over a ground plane

Equivalent Model:

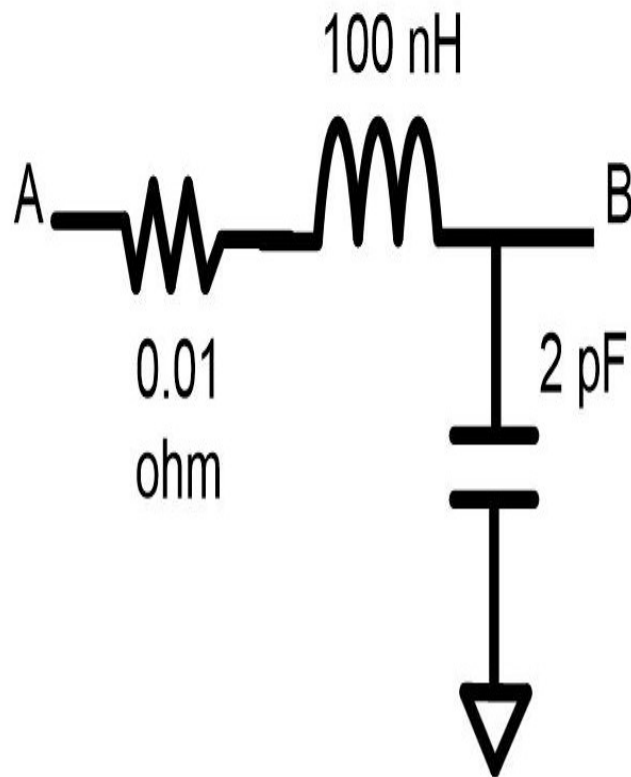
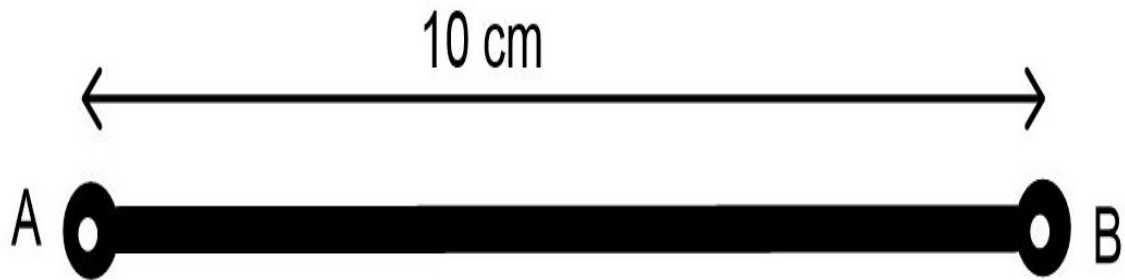


Figure 1-3. Wire Impedance

As shown, a short piece of wire can have significant impedance. A piece of 24 AWG wire, 10cm long, will have 100nH of inductance, 10 milliohms of resistance, and capacitive coupling to the environment around it. In this example, the wire impedance and capacitance to the outside environment will act as a LPF above 300-400 MHz, and be sensitive to motion, due to capacitance changing as a function of wire placement relative to ground. Larger wire can reduce resistive losses and parallel wire connections can reduce inductance some, but the inductance is not easily removed.



- PCB Connection
- 0.5 Oz copper thickness
- 3 mm wide trace
- Placed 0.5mm over a ground plane

Equivalent Model:

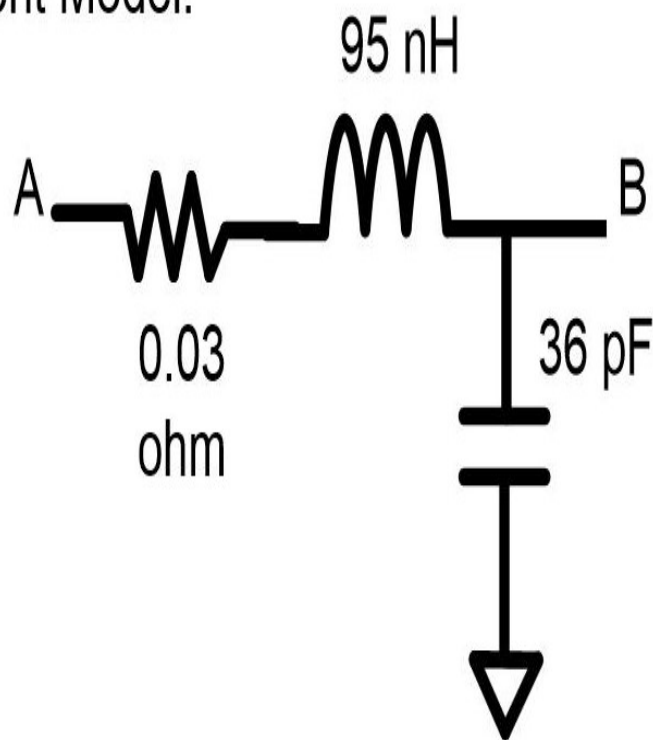


Figure 1-4. PCB Connection Impedance

A similar 10cm connection on a PCB shows identical characteristics. In this example, the inductance is about the same as the wire, and the capacitance has increased due to the connection being tightly spaced over a ground plane. For this situation, the PCB trace will start to perform as a LPF around 80-90 MHz due to the increased capacitance. One advantage of the PCB trace over the wire is that the characteristics of the impedance will not change due to the fixed environment that the PCB creates.

The takeaway here is that every connection has impedance, and some coupling to an outside environment.

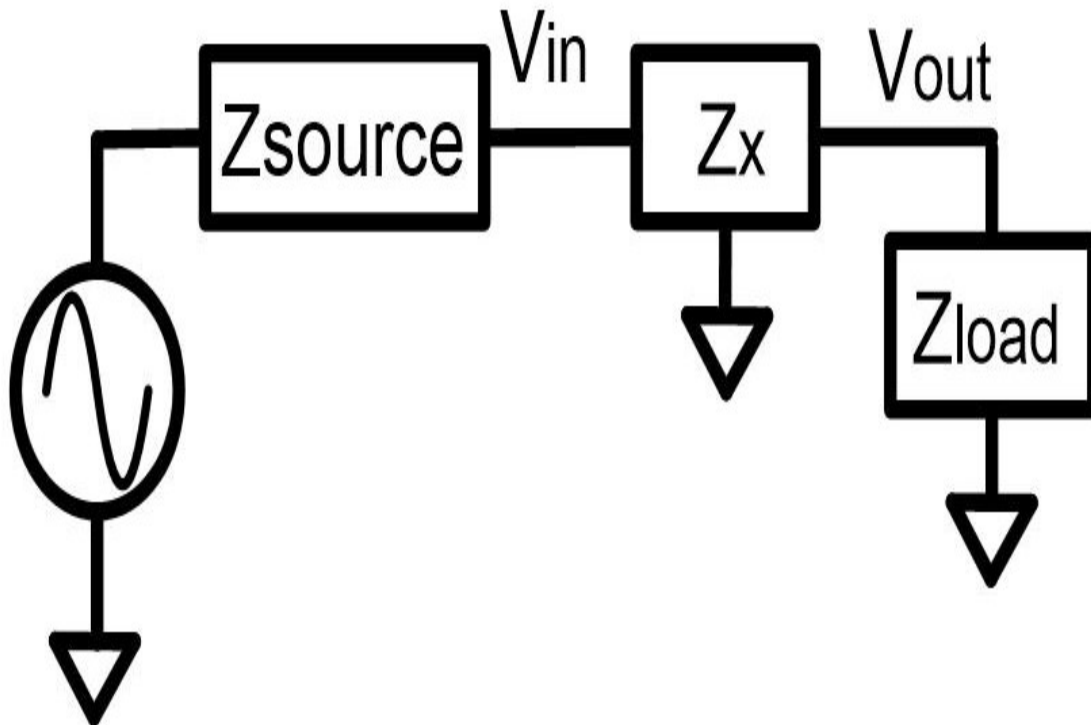


Figure 1-5. Connection Impedance Voltage Division

When connecting things together, the impedance of the source, the connection, and the load, all play a part in how much signal loss occurs. Signal loss and distortion get progressively worse with both longer connections and higher frequencies. Designers working under 50 MHz and

on a small (10cm x 10cm) PCB generally can ignore much of this and survive. High bandwidth, long distances, and connections off the PCB make interconnect issues a significant part of the design problem.

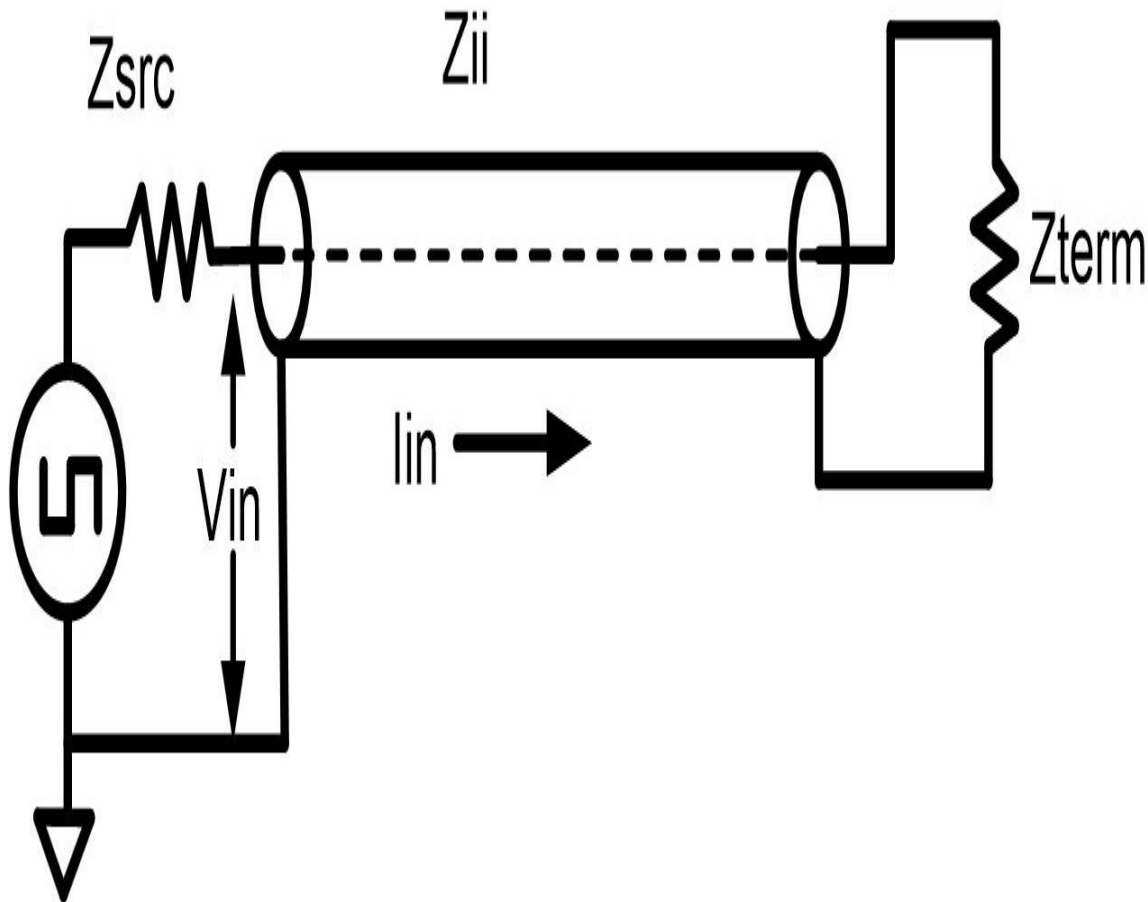


Figure 1-6. Transmission Line

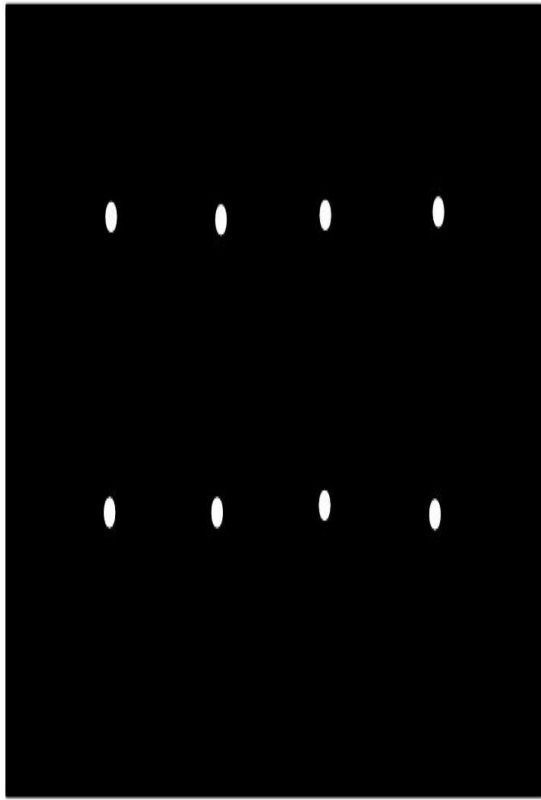
In addition to loading and losses associated with connection impedance, high frequency signals with a lengthy connection also exhibit transmission line characteristics. For the example of the 10cm connection, current takes about 0.7 ns to transit the wire. Depending on connection length and the frequency of the signals involved, improperly terminated transmission lines can affect signal integrity.

Consequently, designers need to consider transmission line issues when the signal wavelength becomes a significant part of the connection distance.

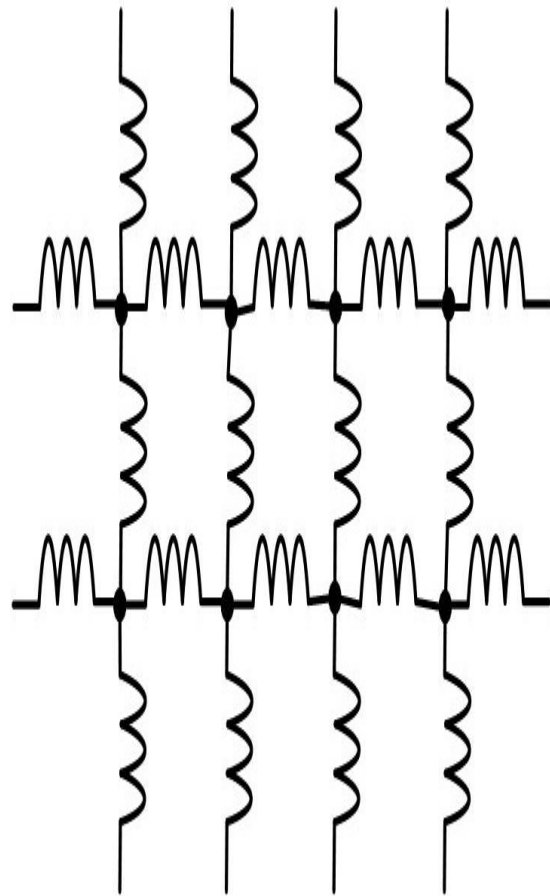
For the 10cm wire example, and keeping under 0.1 wavelength on the wire, would limit this at a 280 MHz sinusoid. A digital signal has multiple harmonics that need to be included, further limiting the connection's capability.

A termination impedance (Z_{term}) should match the characteristic input impedance (Z_{ii}) of the line to minimize reflections on the transmission lined. Practical applications terminate at both ends of the line (Z_{src} and Z_{term}) to minimize both the initial reflection, and any residual reflection that that went back to the driven end.

The use of impedance matching, striplines and creating data paths with high signal integrity are topics covered in the chapters on data communication and PCB layout methods.



Solid Copper Ground Plane



Equivalent Model

Figure 1-7. Planar PCB Connections

Even a solid copper sheet used as a PCB ground plane will exhibit resistance and inductance between locations. A current surge anywhere on a dedicated ground plane will cause that location to have a voltage transition relative to other locations on the ground plane. A surge current into the ground is analogous to a person bouncing on a trampoline. That ground bounce can be kept low by minimizing ground impedance and minimizing the magnitude of current surges into the ground.

It is important to recognize the concept that there is not a singular ideal ground, but rather that the voltage of ground can vary, with both proximity

between points and the characteristics of the current dynamics passing through the ground plane.

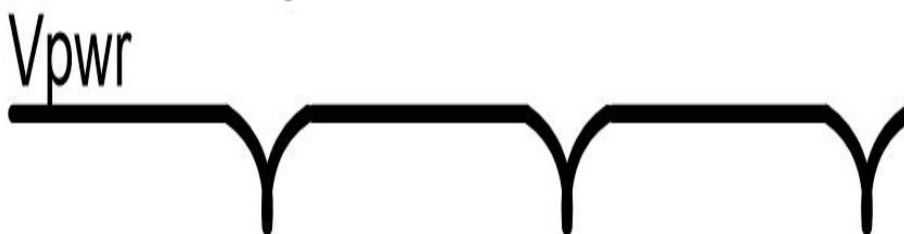
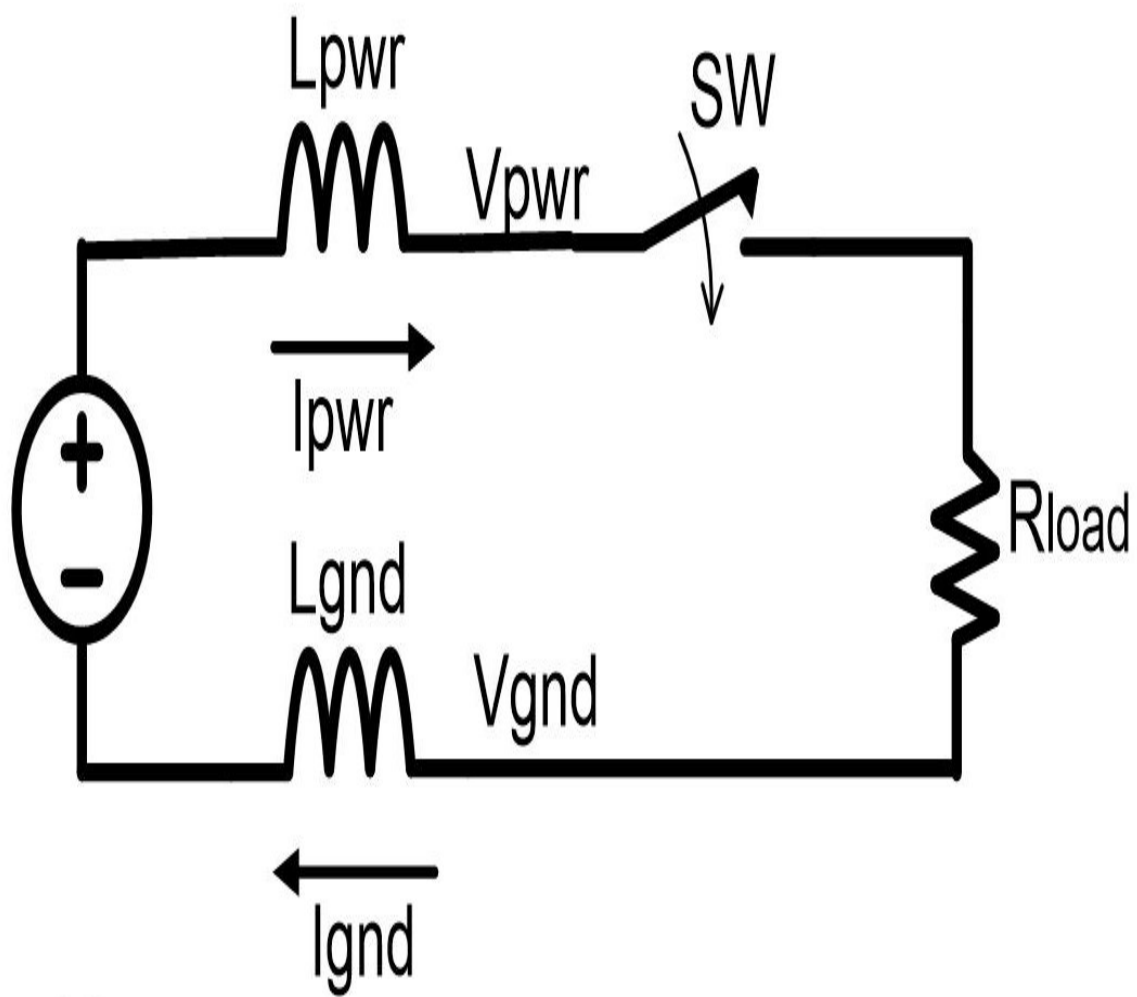


Figure 1-8. Power & Ground (P&G) Impedance, Ground Bounce and Power Instability

Both power and ground connections exhibit impedance. The typical electronic device presents switched loads between power and ground, causing the power voltage to drop and the ground to rise. The magnitude of P&G variance is influenced by the connection impedance and the transient rate at which the current changes (dI/dt). The transient current can be changed through the load resistance variance, and how much capacitance is present across the load. Proper system design adjusts these variables to give good enough P&G stability to maintain proper functionality. In most situations P&G bounce can be kept manageable with proper circuit techniques, power bypass filtering and PCB layout.

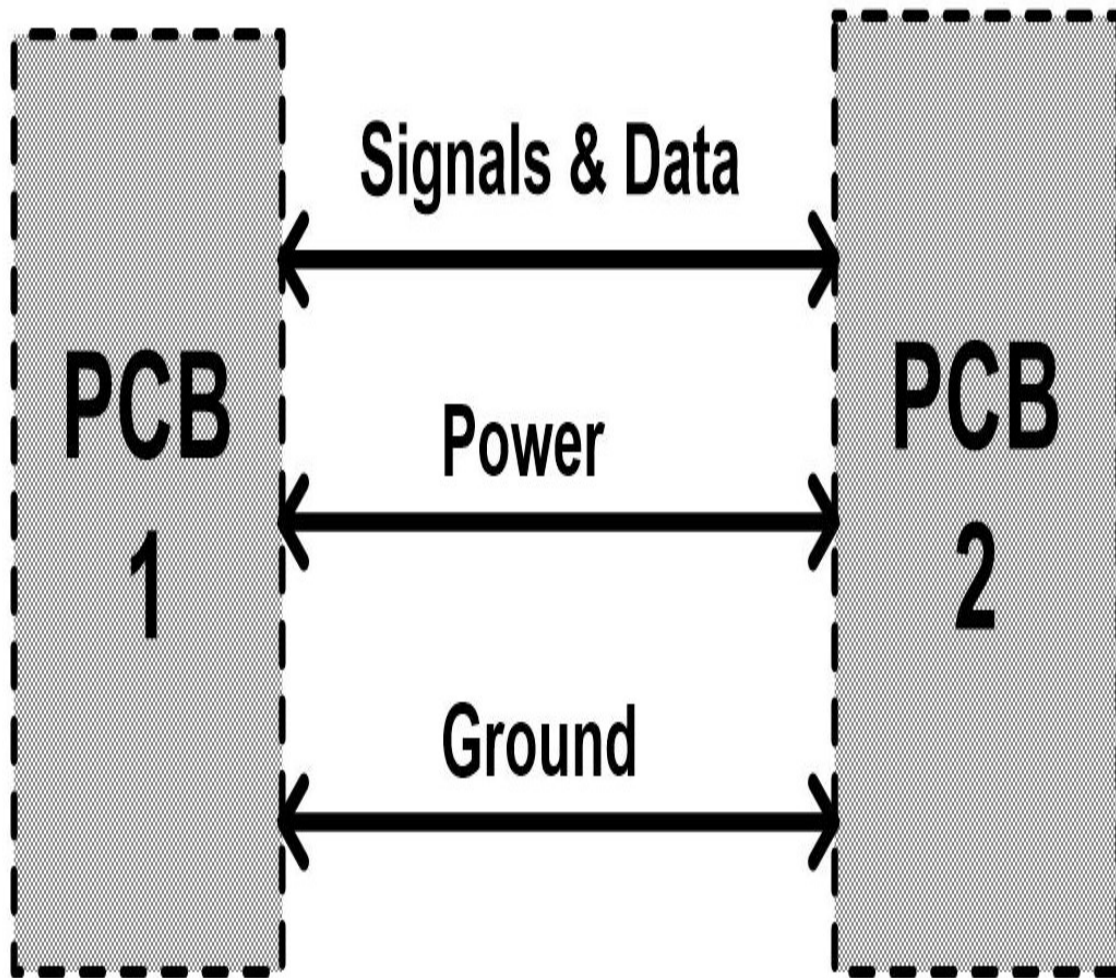


Figure 1-9. Impedance Between Circuit Boards

Connections between circuit boards will have signal integrity problems on all connections. Current surges on P&G wires will create voltage variances between the separated boards. That variance makes noise on any ground referenced signal, relative to the local ground. External stimulus like EMI or electrostatic discharge (ESD) can further corrupt all signals. Special efforts need to be made to “clean up” the noisy raw power coming into the board and create clean power on the PCB. Also, data and signals need to be passed between boards in a manner that is not dependent on local grounds or power supplies. The chapters on power systems, noise, and digital communication deal with these issues.

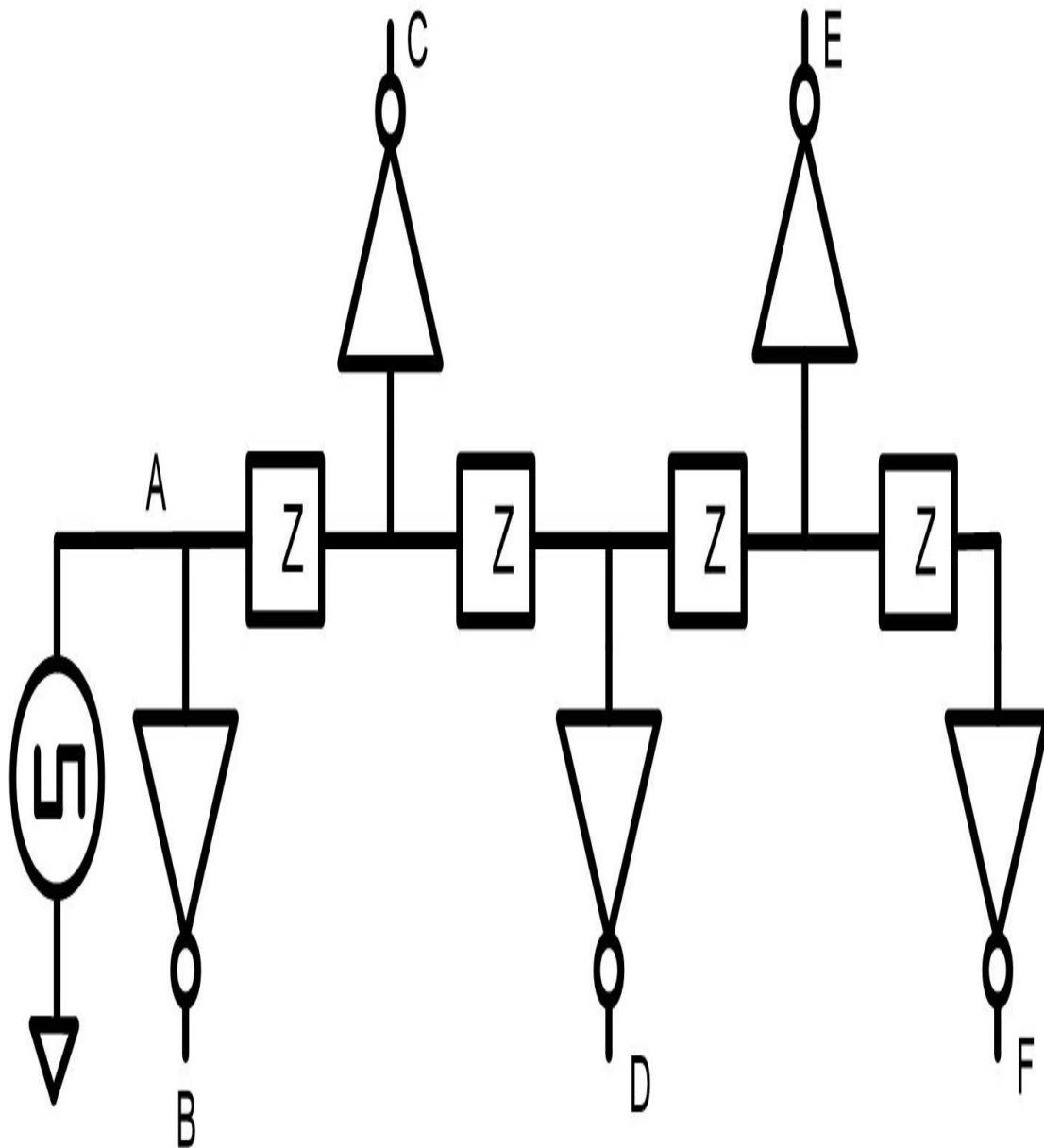


Figure 1-10. Phase Errors in Digital Signals

Distributing a common signal to multiple receivers can result in phase errors between signals received. The figure shows one signal going to five locations. The distributed connection impedance, the path length, and the capacitive receiver loading, all results in five different phases of the original

signal. This is a common problem in “clock tree distribution” especially with high clock frequencies, multiple destinations, and longer distances.

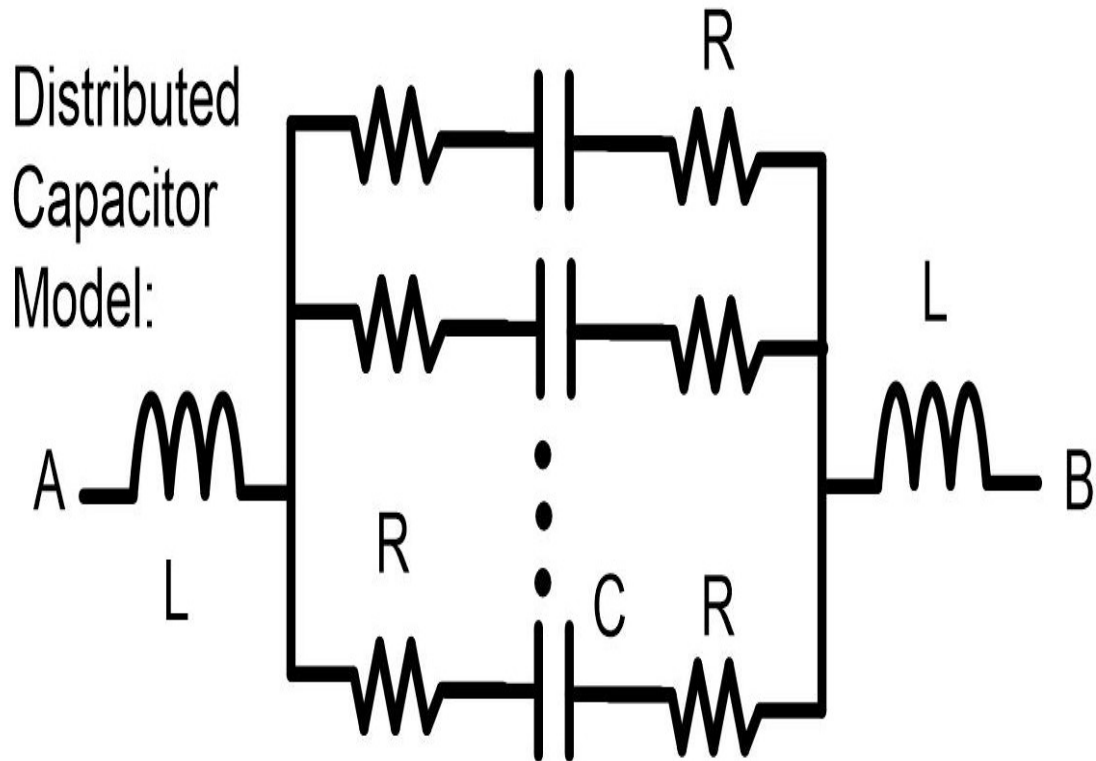
Frequently, designers don’t consider interconnect issues until problems arise. At that point, it’s an expensive redesign. When high currents, long connection paths or high frequencies are involved it needs to be carefully considered in the design.

Basic Components

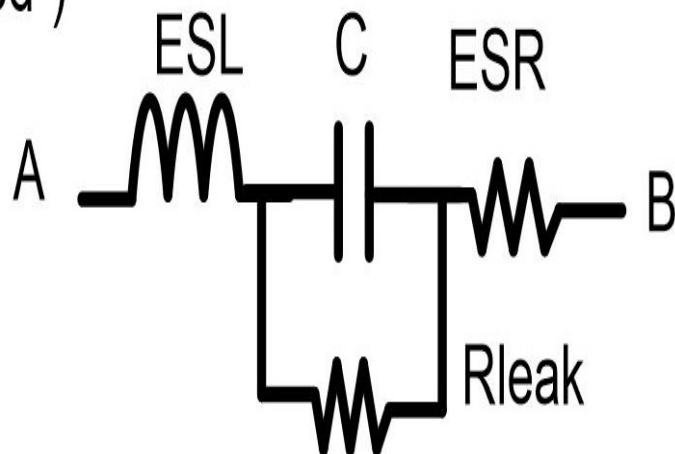
The evolution of electronics includes a diverse collection of devices many of which are now obsolete. The emphasis here is on modern systems that will be commercially manufactured in volume. Axial lead devices and other through-hole components are minimized in most high-volume manufactured devices. Consequently, surface mount circuit components are the focus here.

Capacitors

A distributed element model of a “Surface Mount Multi-Layer Ceramic Capacitor” (SMT MLCC) is shown.



Equivalent
("Lumped")
Model:



$$C = f(V_{AB}, \text{Temperature})$$

Figure 1-11. Surface Mount Multi Layer Ceramic Capacitor

This type of capacitor has multiple interleaved conducting and insulating layers. The actual capacitance is between adjacent plates (C), all plates have a small amount of resistance (R), and the interconnect has inductance (L). A distributed element model can be awkward to work with, and an equivalent model is generally sufficiently accurate.

The equivalent model, also known as a “lumped element” model includes elements which model externally observed behavior. This device includes a single ESL “Equivalent Series Inductance (L)” ESR “Equivalent Series Resistance”, a single capacitor and a leakage resistance. The principle capacitor element can vary, primarily as a function of temperature and the voltage across it. The variability is due to the insulating dielectric characteristics used between the plates.

Depending upon application, some elements can be ignored, or in other scenarios, limit the device performance. ESL is important for radio frequency (RF) circuit performance and high frequency power filters, but does not affect low frequency performance. ESL with the C creates a self resonance tuned circuit, which limits the high frequency response of the capacitor.

1uF SMT 0603 Cap, Self Resonance @ 5.5 MHz

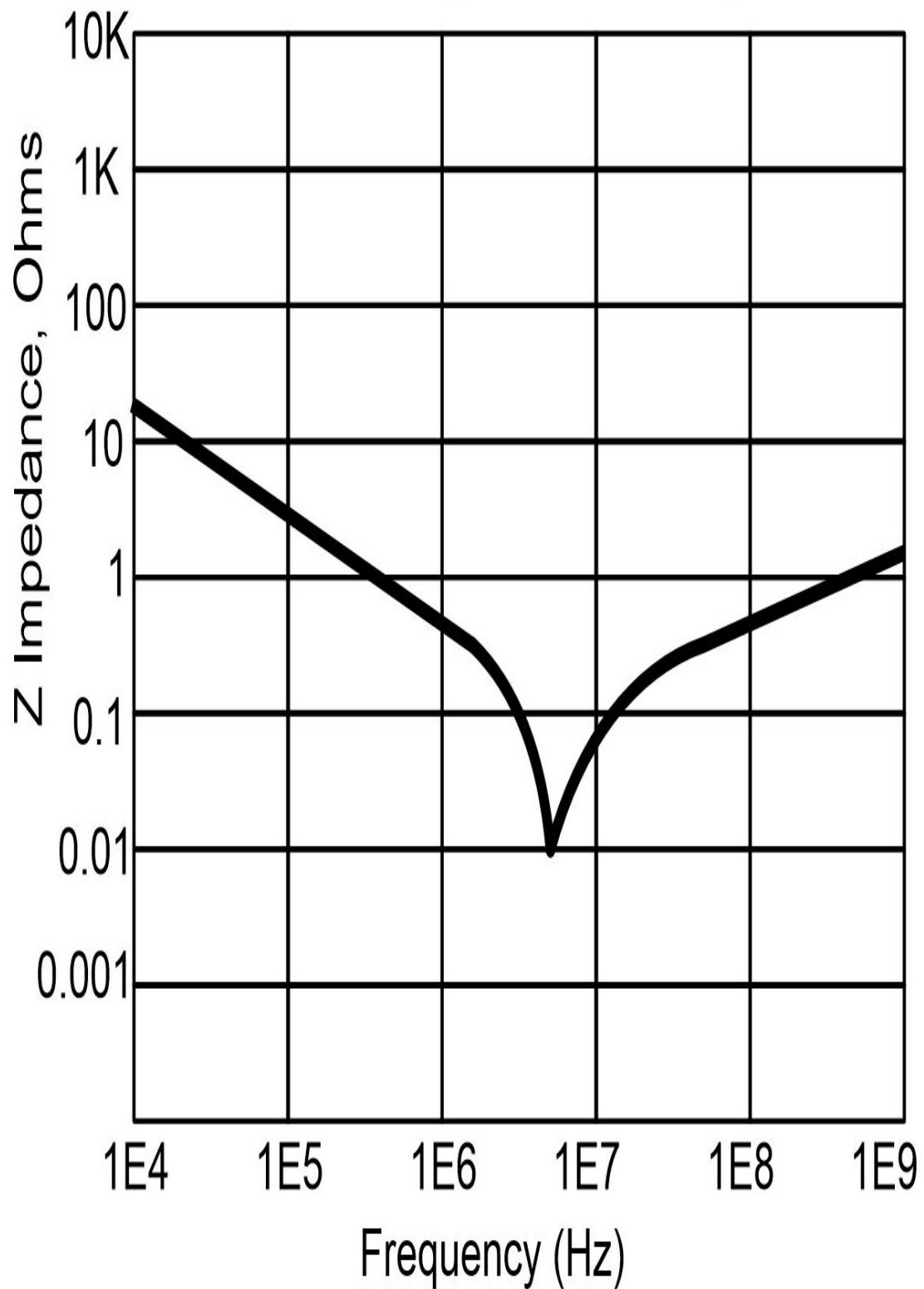


Figure 1-12. Self Resonance of SMT MLCC (placeholder)

Leakage resistance becomes an issue if a capacitor is used to hold a static charge for an extended time. ESR becomes apparent in circuits using high surge currents through the capacitor.

The SMT-MLCC makes up the majority of PCB capacitors due to excellent reliability, low cost, and wide selection options available.

Modern capacitors make tradeoffs between:

- capacitance per unit volume
- maximum applied voltage (breakdown voltage)
- min-max operating temperatures
- temperature variation from nominal value
- applied voltage variation from nominal value
- aging variation from nominal value
- nominal value accuracy
- total overall life

These parameters can interact, if the manufacturer changes one parameter it can affect another. One example: Creating a higher breakdown voltage implies a larger body size for the same value of capacitance.

SMT-MLCC capacitors have different dielectrics, which leads to different performance parameters. Looking at the more common devices, some brief comments are useful to better understand the differences:

COG, NP0:

These are known as Class 1 capacitors, designed for minimal thermal variance, and minimal change from voltage bias. Class 1 devices sacrifice capacitance per unit volume in order to achieve higher accuracy and stability. For a given body size, the total capacitance is

limited. These devices are useful for tuned circuits and other applications, where accuracy and stability are needed.

X5R, X7R, Y5V, et al:

These are known as Class 2 capacitors, their design sacrifices voltage bias accuracy and thermal stability to achieve more capacitance within a given volume. The notation used for these devices is a little confusing, but actually simple to decode. The three characters decode to a minimum temperature, maximum temperature, and value variance over temperature.

Class 2 Ceramic Capacitors Rating Code:
Temperature vs. Tolerance

Low Temperature	High Temperature	Temperature Variance
X = -55 degrees C	4 = 65 degrees C	P = +/- 10%
Y = - 30	5 = 85	R = +/- 15%
Z = +10	6 = 105	S = +/- 22%
	7 = 125	T = +22%, -33%
	8 = 150	U = +22%, -56%
	9 = 200	V = +22%, -82%

From: EIA RS-198

Figure 1-13. EIA RS-198 Code for Class 2 Capacitors

An X7R capacitor is designed for use from -55C to +125C and will have +/- 15% variance over temperature. The Y5V capacitor is designed for use from -30C to 85C and will have from +22% to -82% in value. Picking the proper capacitor dielectric can significantly affect accuracy.

Class 2 devices also exhibit “DC Bias Effect” also known as the “DC Voltage Characteristic” which consists of changes to the capacitance value as a function of the static DC voltage across the device. Generally, the value of the capacitance decreases as the bias voltage increases. This can be significant, with as much as 60% of the capacitance value changing, depending on the specifics of the device. Devices with a higher breakdown voltage tend to have less DC Bias effect for the same voltage change, if a designer needs to reduce this effect. Need to check your specific capacitors and determine how much this affects your design.

Entire books have been dedicated to capacitors. A closer look at the limitations of the SMT-MLCC is warranted because it is the predominant capacitor used in modern designs.

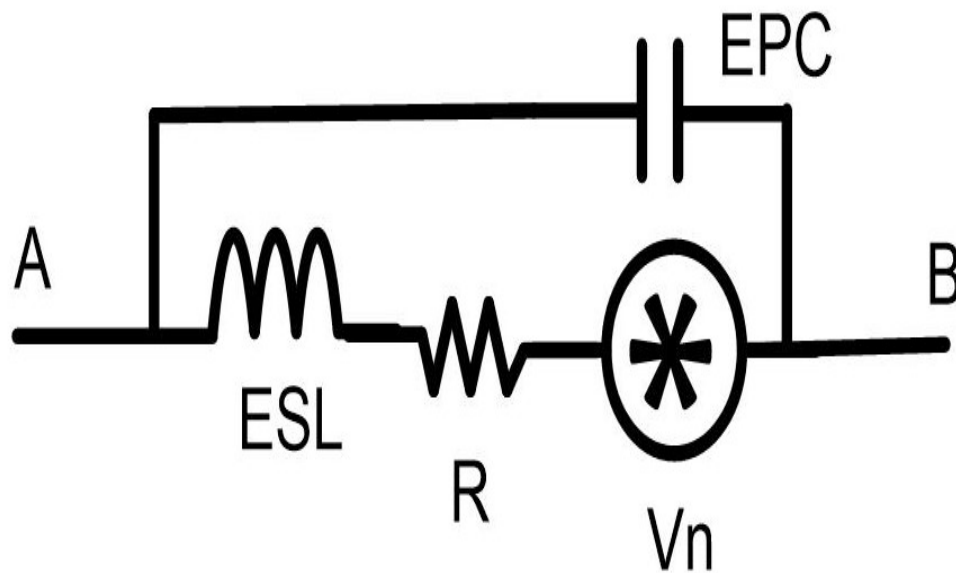
Several other capacitors are frequently used in modern designs. The Aluminum Electrolytic Capacitor (AEC) is widely used in DC power supply filters and other applications where large amounts of capacitance in a small package, and low cost, are needed. All AECs suffer from poor high frequency response due to high ESL, so AEC's are not suitable for RF situations. High frequency response can be supplemented by parallel SMT-MLCC devices if needed. The AEC also has a limited life, fussy temperature restrictions, and significant ESR. Check device specifications for ageing information and lifetime vs. temperature data. Most AEC devices have performance and lifetime parameters that are unique to a specific manufacturer and product line.

Both tantalum and aluminum-polymer capacitors are available in high reliability and long life variants. Component selection here needs to be done on a case by case basis because these devices also have short life variants as well.

Important things to look for in a capacitor:

- Nominal component value and fabrication tolerance
- Breakdown voltage
- Package size
- Dielectric type, temperature range, variance over temperature
- DC Bias Effect and device variation due to bias voltage
- ESL characteristics and self resonance frequency for high frequency applications
- ESR characteristics where series resistance can affect performance
- Reliability, lifetime and aging data

Resistors



$$V_n = \sqrt{4 K_b T R \Delta F}$$

V_n = RMS Noise Voltage

K_b = 1.38×10^{-23} Joules/Kelvin

ΔF = Bandwidth, Hz

R = Resistance, Ohms

T = Temperature, degrees Kelvin

Figure 1-14. -x: High Frequency Resistor Model

A typical resistor model is shown in **Figure 1-29**-. In addition to the ideal resistance “R” some additional elements need consideration. These are: Equivalent Series Inductance (ESL), Equivalent Parallel Capacitance (EPC), and internally generated noise (V_n) created by the resistance itself.

Ideally a resistance creates a flat impedance over frequency. Because of EPC and ESL, the impedance may not be flat at high frequencies. For DC bias applications, the ESL and EPC can be ignored. Multi watt resistors are commonly fabricated with wire wound methods, and have significant ESL and EPC components.

Most modern surface mount resistors are some type of film resistor: thick film, thin film, and metal film are common descriptions of devices available. Multiple vendors have found many different approaches to making resistors. Generally, metal film resistors are lower noise than carbon and thick film devices, but check specific vendors for noise data.

Thermal noise, also known as Johnson-Nyquist noise is created by the thermal agitation of electrons in a resistive material. For most large signal applications this noise is not a significant issue. It does become important in RF front ends, communication channels, and low amplitude scenarios, where signals are in the nanovolts.

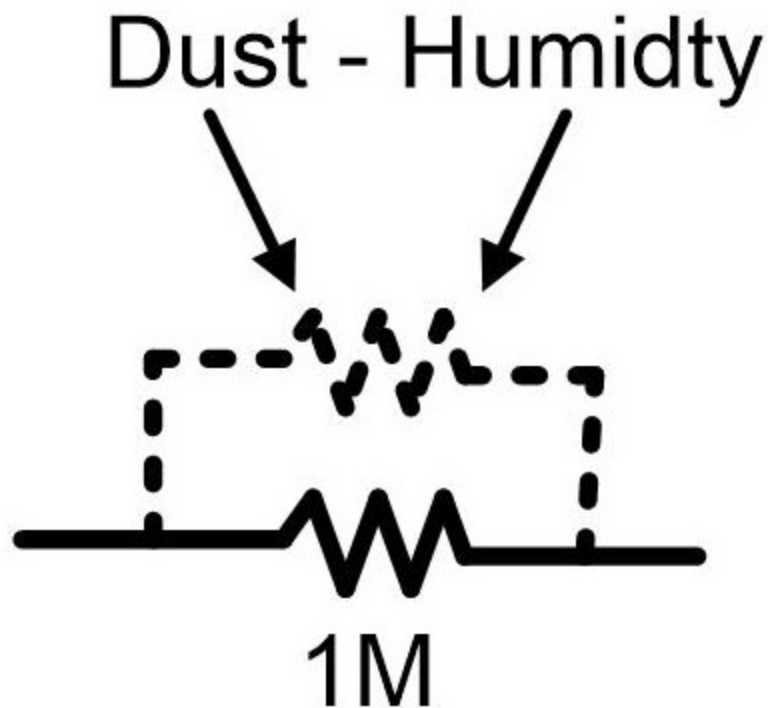


Figure 1-15. Environmental Sensitivity of High Impedance Resistors

Gigaohm resistors are available from manufacturers, but special considerations have to be made to use them. Dust and humidity in the environment can have a lower impedance path than the resistor. As a general rule, anything over 100K ohms needs special environmental consideration to avoid alternative current paths around the resistor. Most modern designs don't need high value resistors, because most analog circuits that used them in the past have been replaced with more reliable digital methods.

Important things to look for in a resistor:

- Nominal component value and fabrication tolerance
- Material Composition, film resistors are preferred, older carbon based resistors should be avoided
- Power rating

- Temperature Range, min-max
- Maximum voltage rating, for high voltage applications.
- Thermal variance, sometime specified as Temperature Coefficient of Resistance
- ESL for high frequency situations
- Thermal Noise, Carbon Contact noise
- Stability, repeatability, aging

Inductors

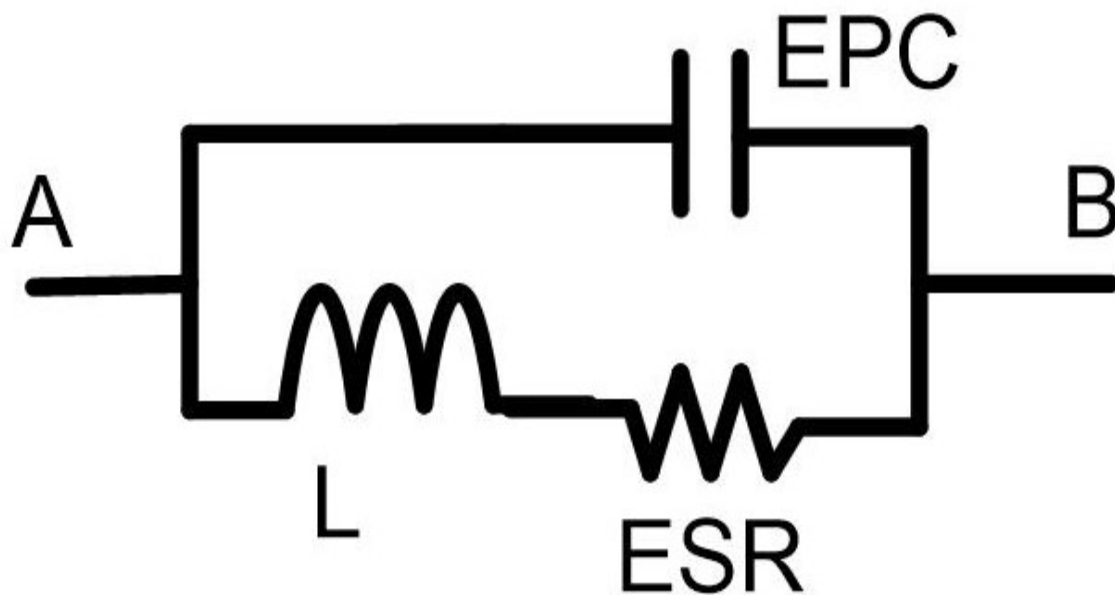


Figure 1-16. -x: High Frequency Inductor Model

A high frequency model of an inductor is shown in **Figure 1-29**. Some Equivalent Parallel Capacitance (EPC) exists, which limits the useful frequency range. With EPC and L in parallel, the device will resonate as a tank circuit and look like an open circuit at the Self Resonance Frequency (SRF). Below SRF the device functions as an inductor, above SRF and the device performs like a capacitor.

The Equivalent Series Resistance (ESR) defines the quality factor (Q) of the inductor when used as part of a filter. When used in a switching power converter, ESR will limit power efficiency.

Inductors are commonly used in power conversion, RF signal processing, and EMI limiting. Inductors are used in switched mode power conversion, both AC-DC and DC-DC converters. Inductor priorities for power conversion are maximum current, saturation current, thermal range, and resistive losses. Switching converters have high current transients and suitable devices are wire wound on a magnetic core, with high current capability and low resistive losses.

Inductors in RF signal processing are used in tuned circuits and need consistent accuracy for reproducible frequency response. Priorities are high Quality factor, component tolerance/accuracy and making sure the SRF does not affect the signals being processed. Lower currents are used in signal processing inductors and many components exist in surface mount variants.

Inductors used for EMI limiting are commonly known as chokes. The purpose is to pass DC current to reduce/eliminate any transient currents in the connection. High current chokes are commonly done with a wire wound power inductor. Low current chokes can be implemented with smaller surface mount inductors. More discussion of this topic is in the chapter on noise reduction methods.

Important things to consider in an inductor:

- Component value and tolerance, will it meet needed accuracy of the design?
- What is the SRF? Does it affect the signals of interest?
- Current Rating, Peak and Average
- Temperature range
- Core Material and Saturation Current

- DC resistance and device Q
- Shielded or Unshielded

Voltage Sources & Batteries

The ideal voltage source doesn't exist. An infinite source of current at a fixed voltage is an academic concept. Practical implementations of a voltage source can be the output of a voltage regulator, a battery, or another generation source. Invariably, these all have limitations.

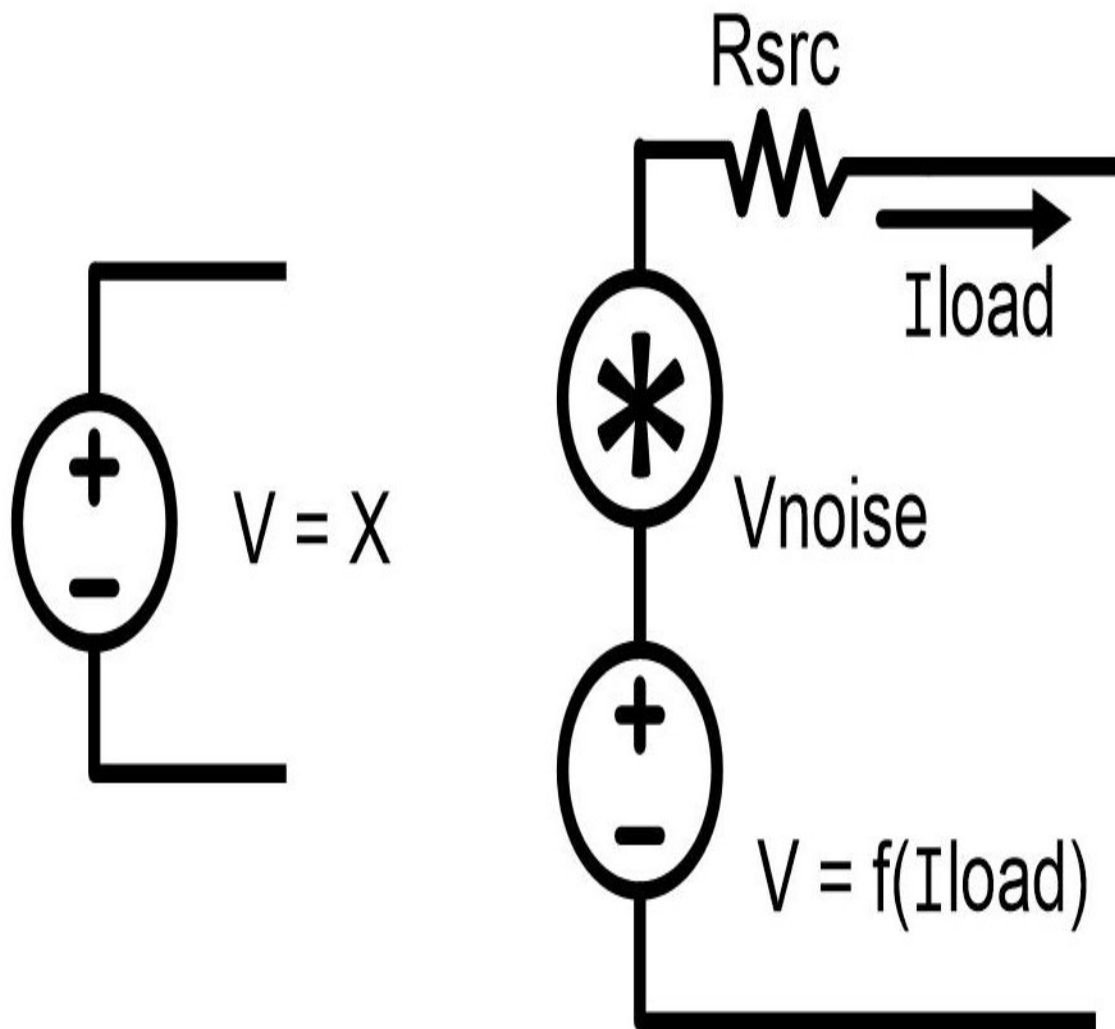


Figure 1-17. -x: Voltage Sources Ideal vs. Real

Any practical voltage source model includes source resistance (R_{srs}), a noise component (V_{noise}) and the voltage source which is a function of the output current (I_{load}). The goal in a design is to minimize the R_{srs} , V_{noise} , and the dependency on I_{load} .

Batteries have a limited life, so voltage variation over discharge has to be added to the model. Also batteries have current limitations on both charging and discharging, reduced performance with ageing, and a myriad of special consideration depending on the particular battery type. The battery systems chapter is devoted to the design, care and feeding of battery supplies.

The characteristics of the V_{noise} signal will be highly dependent upon the methods used to create the voltage source. Linear power supplies and regulators that have been carefully optimized for noise are available when low noise power is needed. Batteries are very low noise but not perfect and can generate electrochemical voltage noise dependent on charge state, or the battery charger can superimpose noise onto the power. Switching power supplies will have high efficiency while generating commutation noise associated with the voltage regulation function. Most modern design work doesn't need ultra low noise power, but any chip designed into a system needs to be investigated for its "noise on the power" limitations.

Current Sources

Similar to the ideal voltage source, an ideal current source doesn't exist. There are a number of circuit designs that come close, with restrictions and limitation on impedance and voltage range.

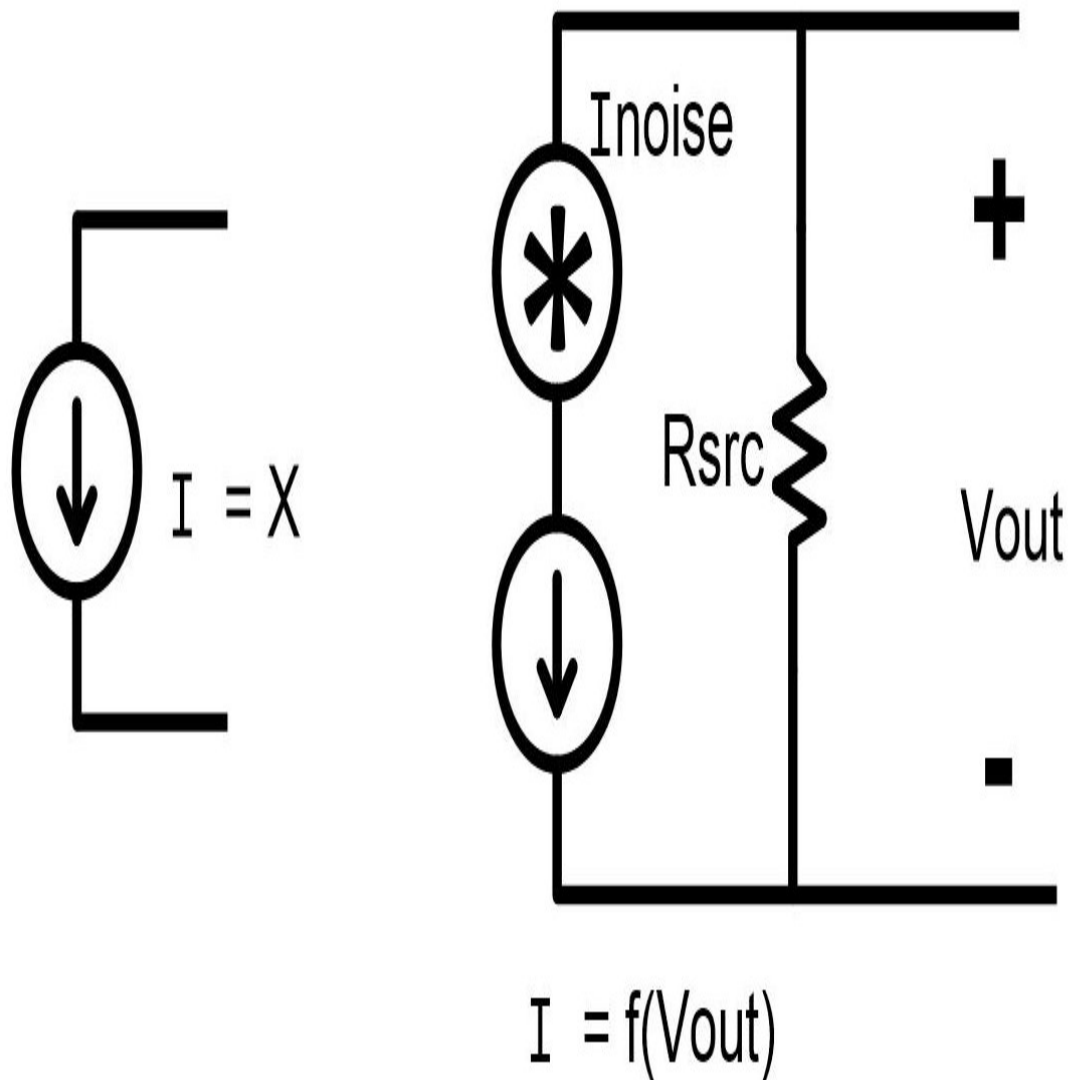


Figure 1-18. Current Sources Ideal vs. Real

A practical current source includes some source resistance (R_{src}), a noise component (I_{noise}) and output current that is a dependent function of the output voltage (V_{out}). Most current sources are limited in the voltage range where it maintains appropriate, fixed current, behavior. Discrete current sources are not commonly used in board level SOC design, but are an important building block within analog and mixed signal IC design. Research the topic “cascode current sources” for further information.

Switches and Relays:

A mechanical switch or relay can be problematic, especially if it is used as a control input on a logic port. Both exhibit similar problems due to their mechanical implementations.

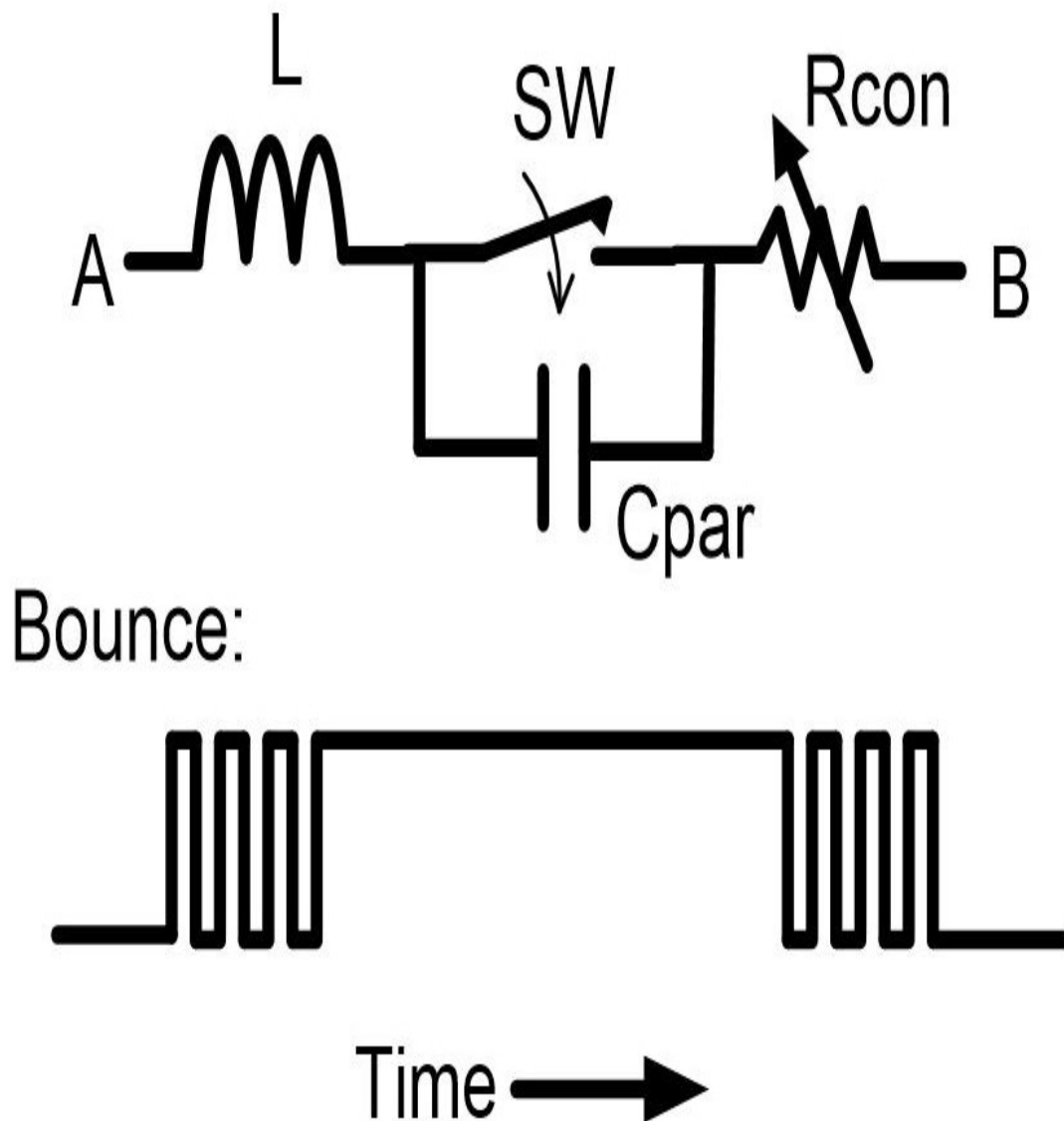


Figure 1-19. Switches with Contact Bounce

Mechanical switches and relays exhibit contact bounce, where the process of opening and shutting contacts creates multiple brief open-shut states. If

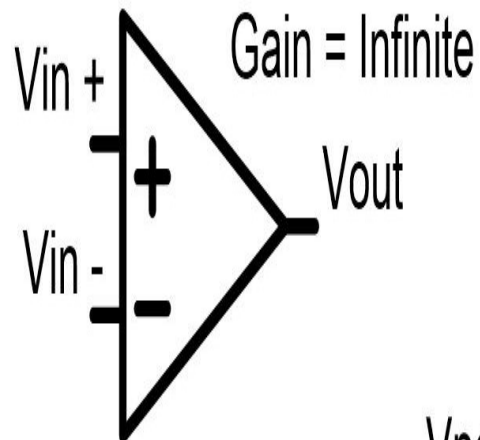
used as a digital logic input, opening and closing contacts are seen as an erratic string of data states due to the imprecise mechanical contacts. Using a switch on a logic port requires software polling the port multiple times and determining state change after remaining stable for 50 -100mS.

In addition to contact bounce, switches have inductance (L) that adds impedance to a closed switch. Parasitic capacitance (C_{par}) allows high frequency signals to pass through an open switch. Also, contacts have highly variable resistance (R_{par}) and can change with every switch cycle. This variability gets worse with age as contacts become dirty or damaged.

Operational Amplifiers

The ideal operational amplifier (op-amp) claims to have infinite gain, infinite bandwidth, infinite voltage range, infinite current output, zero input current, zero noise, zero offset voltage, and, best yet, no power supply is needed. As a math model, it's an interesting idea, but reality is something else.

Ideal:



Typical:

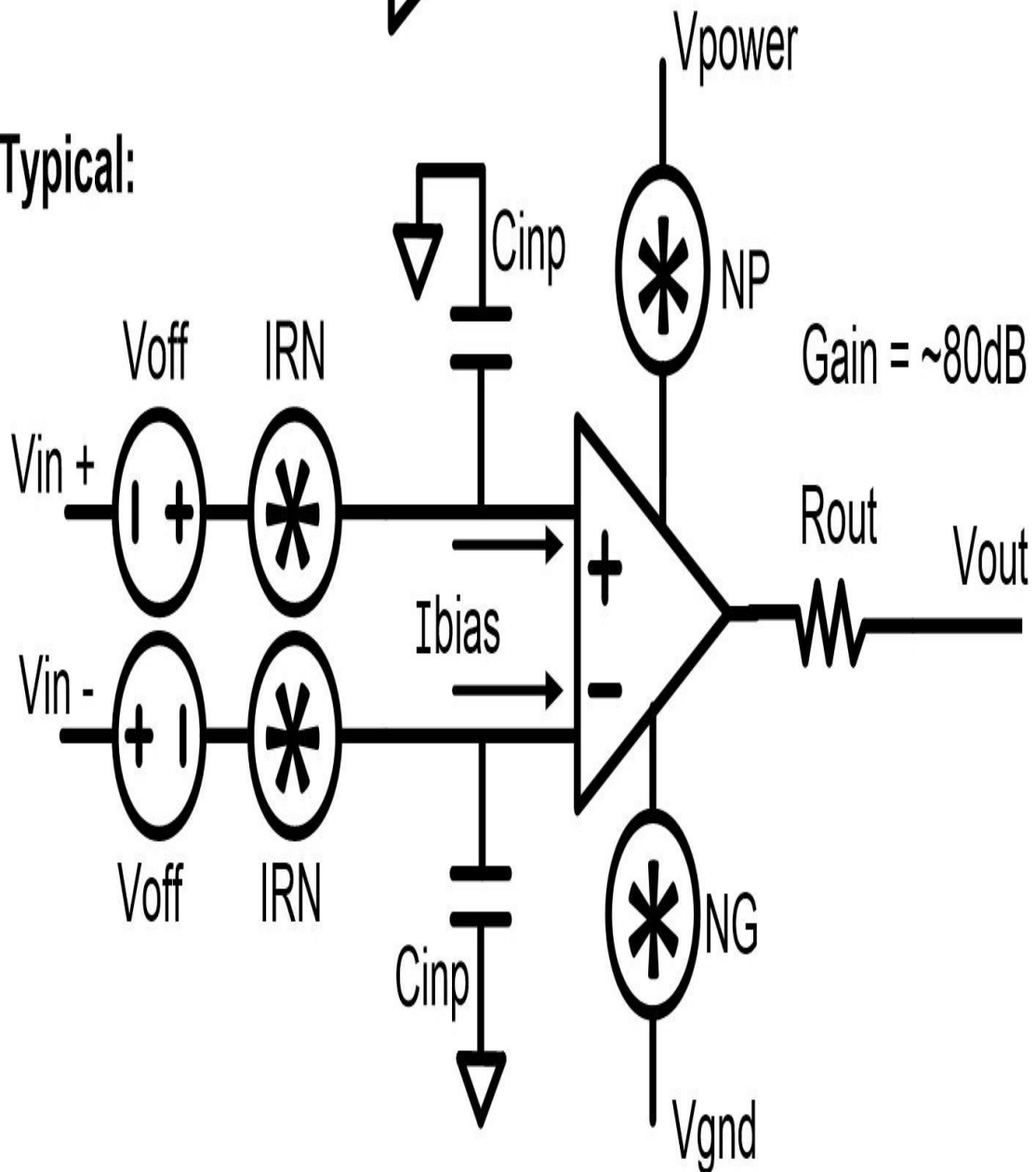


Figure 1-20. -x: Ideal vs. Typical Op-Amp

The real op-amp has limitations that affect performance. First, the device requires power and ground (V_{power} , V_{gnd}) and the output (V_{out}) can be sensitive to noise present (N_P , N_G) on these connections. The output has impedance (R_{out}) and the output voltage (V_{out}) is restricted by the power supply voltage. The gain is not infinite, rather an open loop gain of 80dB is typical.

The gain response over frequency will have both bandwidth limitations and an additive phase response. Also, the output response speed is limited by a maximum “slew rate” at which the device can respond. Since an op-amp is designed to function within a closed feedback system, the high frequency gain of the devices must be internally limited to keep it stable within a feedback configuration. Due to this frequency compensation, op-amps may not be the best device to use in a high frequency design.

Input capacitance (C_{in}) creates loading that can affect high frequency performance. Op-amps designed with bipolar transistors will have input bias current (I_{bias}). The input transistors used within the op-amp, will not be perfectly matched and creates an equivalent input offset voltage (V_{off}) that is typically 1-10mV depending on the specific device. That offset voltage becomes a problem when dealing with small signals or high gain closed loop configurations. The internal circuitry will create noise, which is modeled as an Input Referred Noise (IRN) source.

Many of these limitations are not a problem when working with larger signals, but pushing for high gain, high bandwidth, or using sub-mV signals, will show performance limitations. Performance specifics will depend on the op-amp selected, with many variants commercially available.

Voltage Comparators

A voltage comparator shares many things in common with the op-amp model.

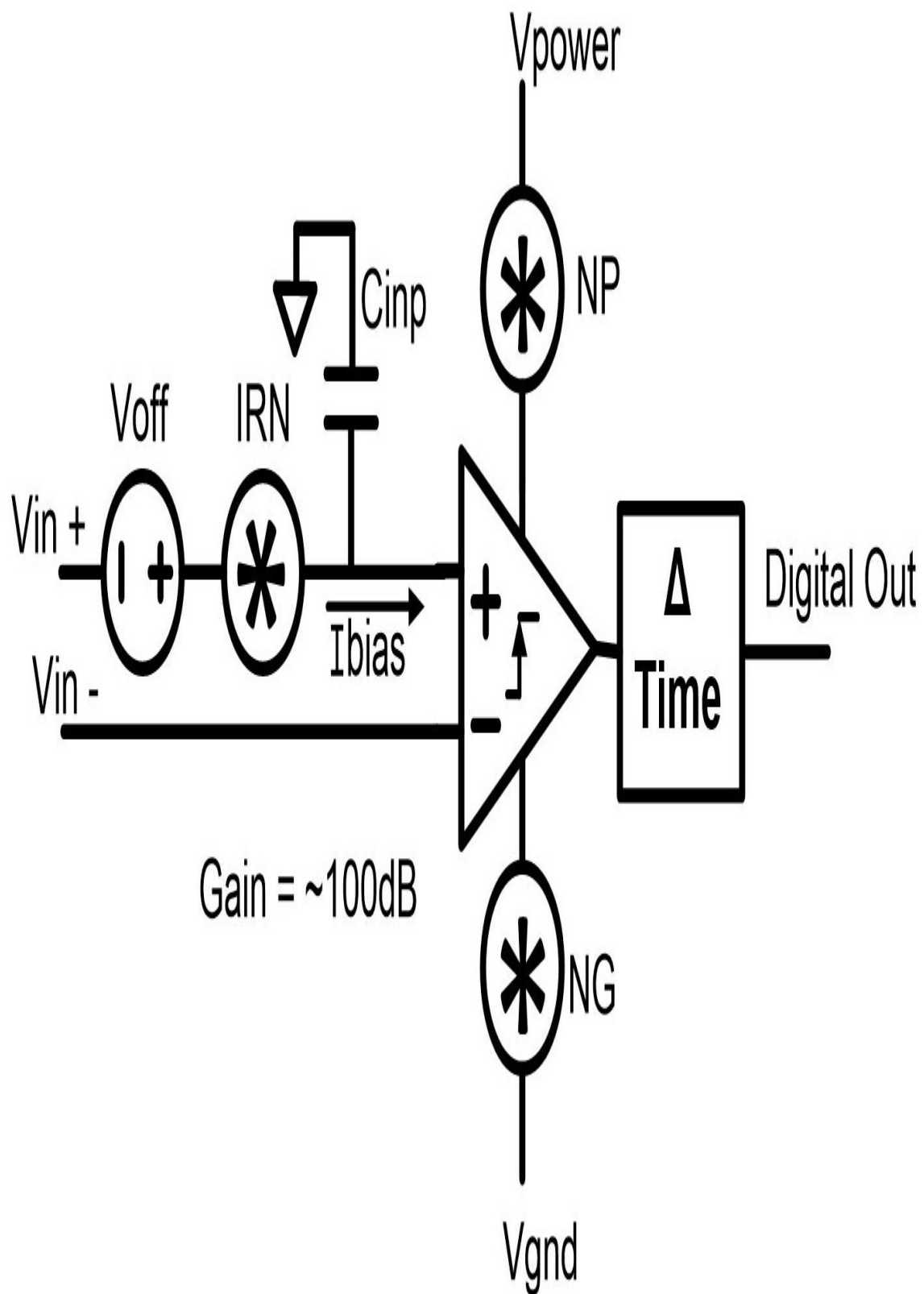


Figure 1-21. Typical Comparator

The comparator also has sensitivity to power and ground noise (NP, NG), input loading capacitance (C_{in}), input referred noise (IRN) and input offset (V_{off}), similar to an op-amp. Since the typical application is against a fixed reference voltage (V_{in-}) a simple one sided model will suffice.

The difference between an op-amp and comparator center is the fact that an op-amp is designed to function within an analog feedback loop and a comparator is not. Comparator gain is often higher (Gain = $\sim 100\text{dB}$), no internal frequency compensation is used, and the output is digital. A variable time response is part of a comparator model since input to output time can be dependent on the input signal amplitude. Small signals with a minimal crossover voltage are slower to propagate through the circuit than a larger voltage transition.

Non Ideal Digital

Digital devices are generally more resilient to noise than analog signals. However, even digital systems have non-ideal problems and limitations.

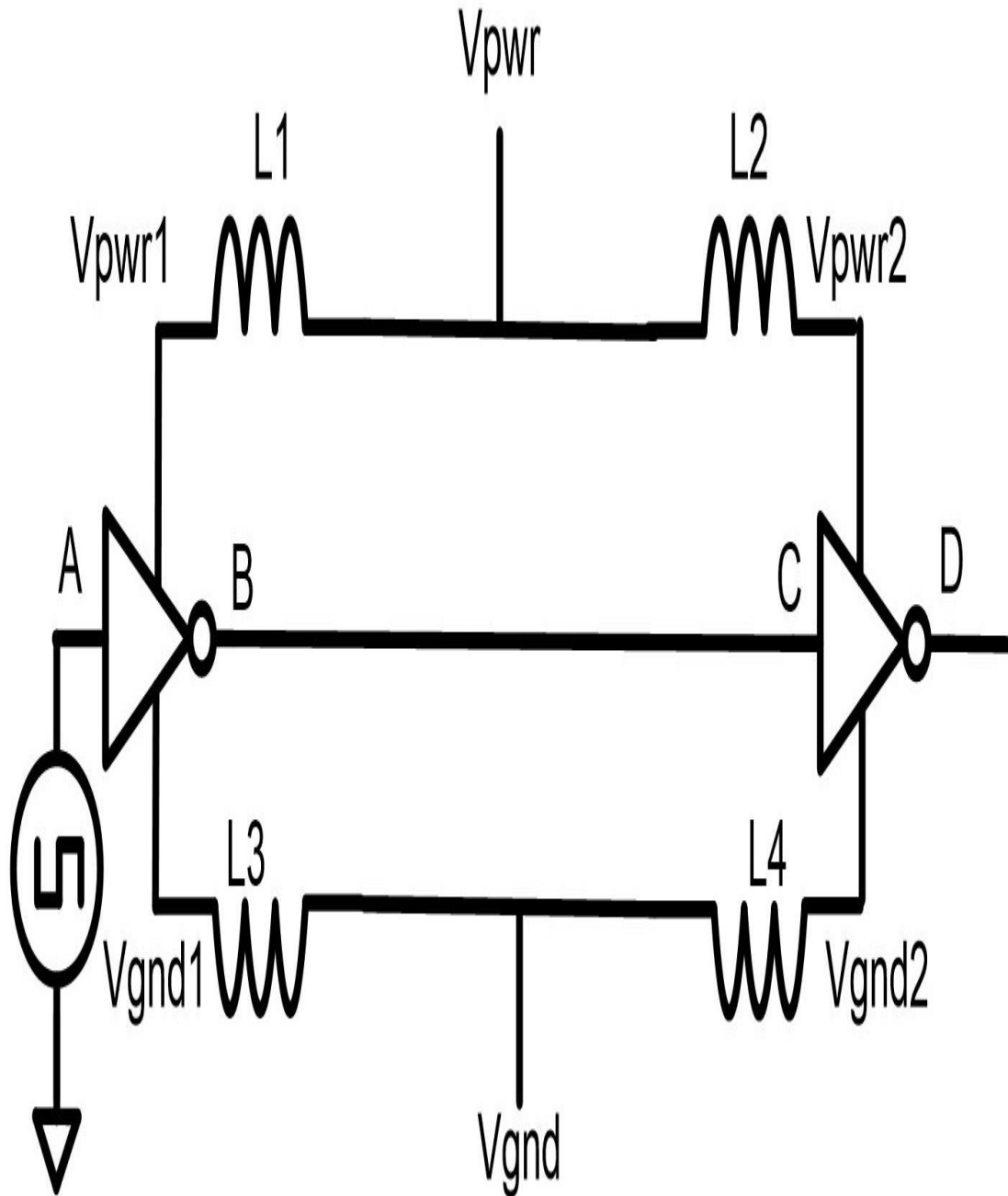


Figure 1-22. Impedance Created Differences Across Power/Ground

Figure 1-29- shows communication between two digital devices. In modern designs, discrete gates are uncommon, but this illustrates communication between any two digital devices, such as FPGA's, micro-controllers, or other devices that share a common power and ground (P&G). As illustrated,

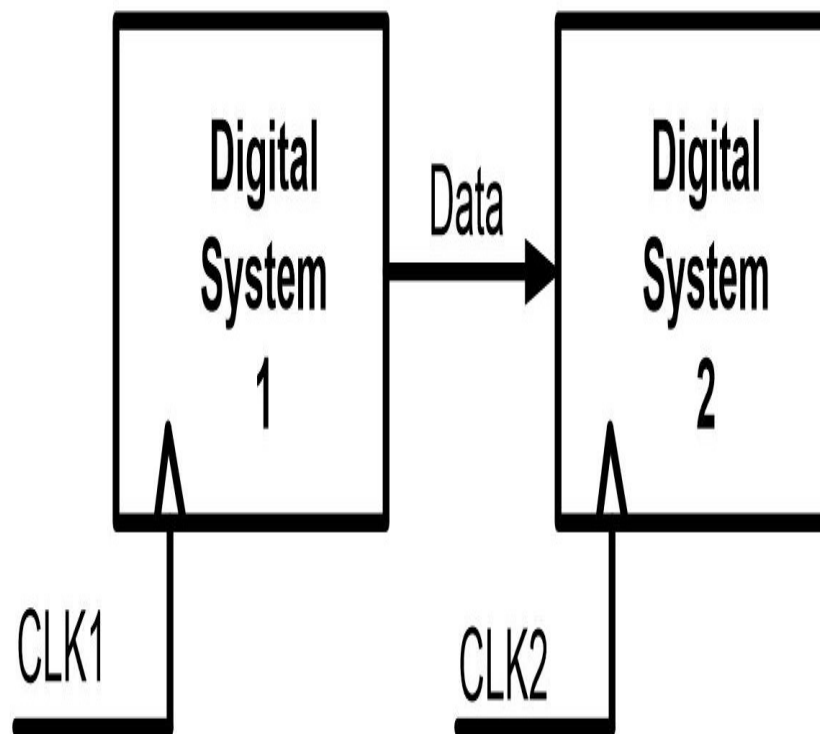
the P&G between devices includes connections with distributed inductance. Depending on what else is connected to the P&G network, currents drawn from, or injected into, the P&G network can cause the connection to reactively dip or rise. As well, the magnitude of that “bounce” can be different at various locations on the network. Illustrated here, the P&G network at one location (PWR1, GND1) has different characteristics than at another location (PWR2, GND2). How much of this bounce is tolerable depends on the transistor technology used in the digital gates, but keeping the P&G stable has to be a priority. Generally, CMOS semiconductors are designed with transistor threshold voltages about one third of the power supply voltage, and can tolerate P&G instability of 20 percent of the nominal power before digital devices start to generate false states. Any commercial product will specify acceptable voltage ranges for both high and low logic states. That specification is a quick guideline to how much P&G noise is tolerable.



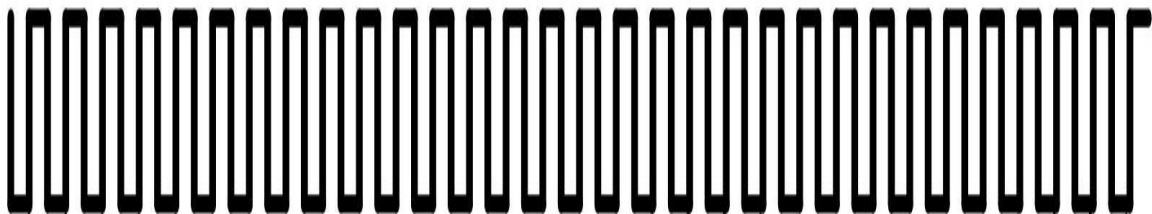
Figure 1-23. Digital Sensitivity to Power and Ground Noise

As shown in the illustration, a clean two state signal at “A” can be corrupted by unstable P&G (PWR1, GND1) at its output (B) – which is then further disturbed by the variation in P&G (PWR2, GND2) at the receiving end (C) leaving an output (D) that no longer represents the original signal.

False state generation is just one of many problems seen in digital systems.



CLK1



CLK2

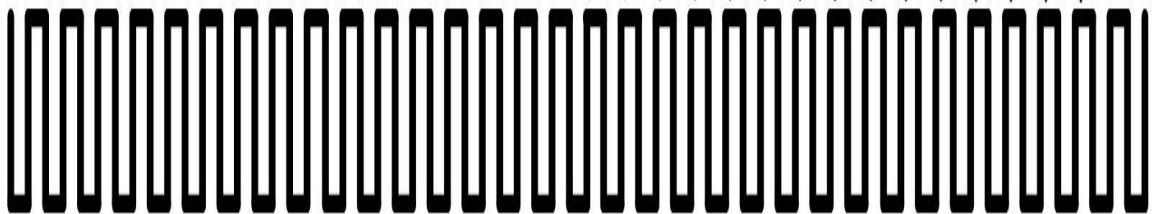


Figure 1-24. Unsynchronized Digital System with Variable Phase

Consider the clocked system shown in **Figure 1-29-** , with two devices communicating at the same frequency but using uncorrelated clocks. Although the clocks are the same frequency, there is no fixed phase relationship between devices. As the phase changes, the data to clock relationship periodically violates setup-hold times and data errors occur. The chapter on data communication covers techniques to deal with reliable non synchronized data transfer.

Asynchronous logic should be avoided due to unpredictable outputs and digital glitches.

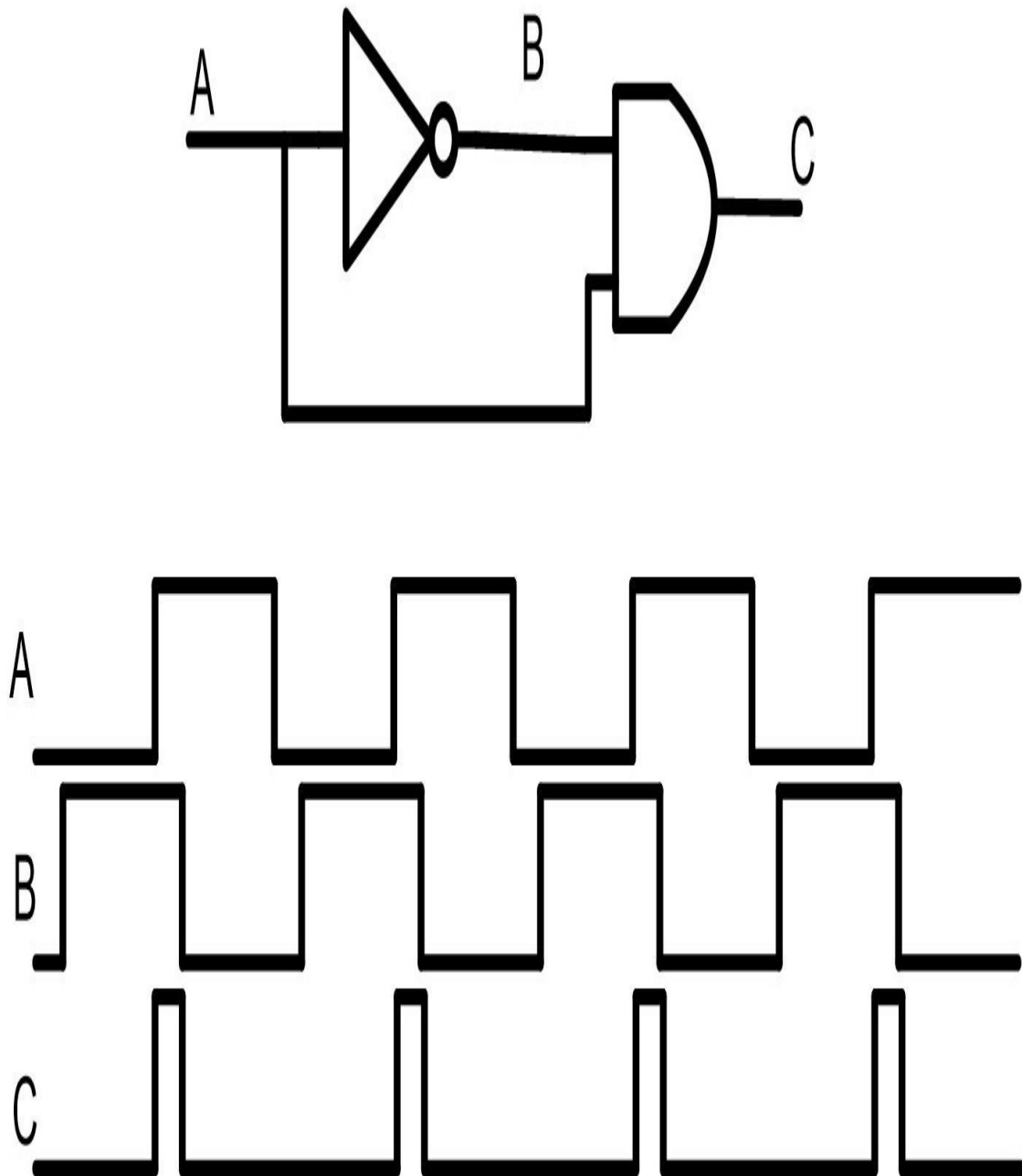


Figure 1-25. Race Conditions in Asynchronous Logic

Although this example is a bit contrived, the Boolean analysis says the output should never assert. Reality shows the propagation delay of the

inverter affects output. This illustrates the motivation for synchronous logic where data resynchronization to a clock is used to avoid glitches.

In addition to problems created by timing relationships, digital functionality can suffer due to the analog characteristics of digital signals. Several factors can contribute to this: higher data rates, resistance of the transmission path, or a transmission path that is loaded with capacitance.

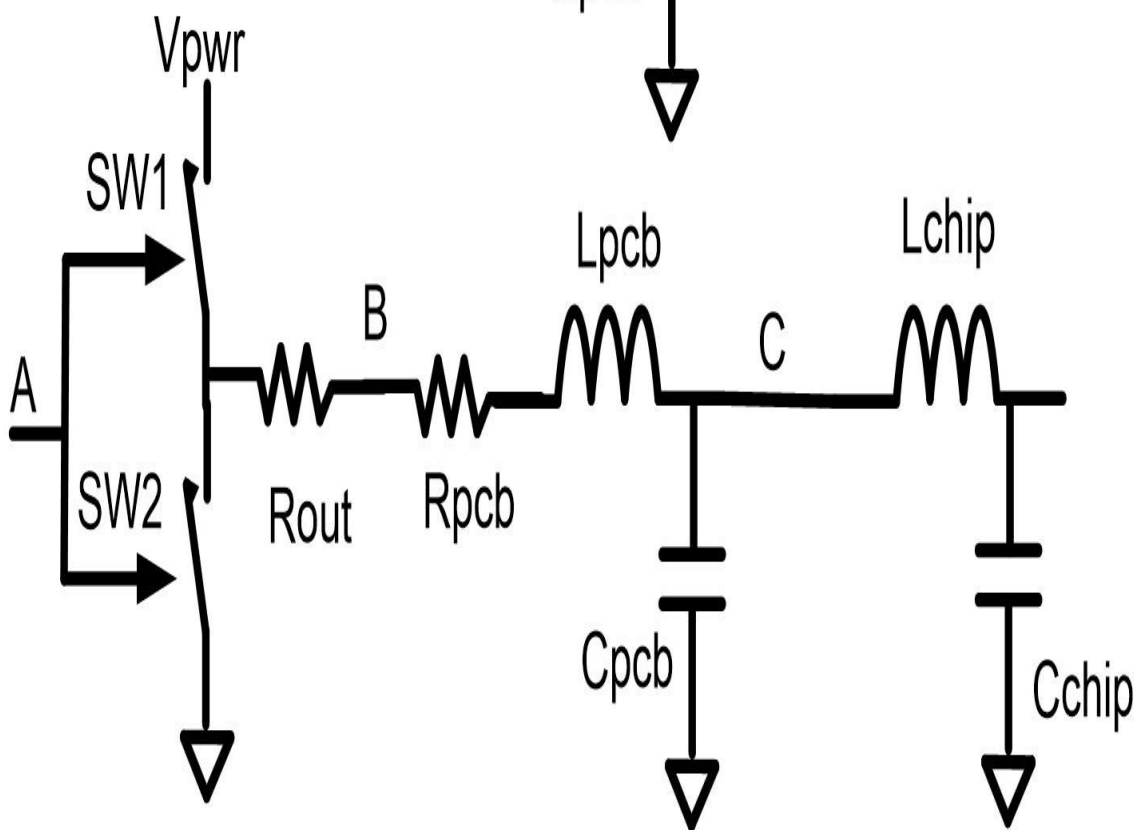
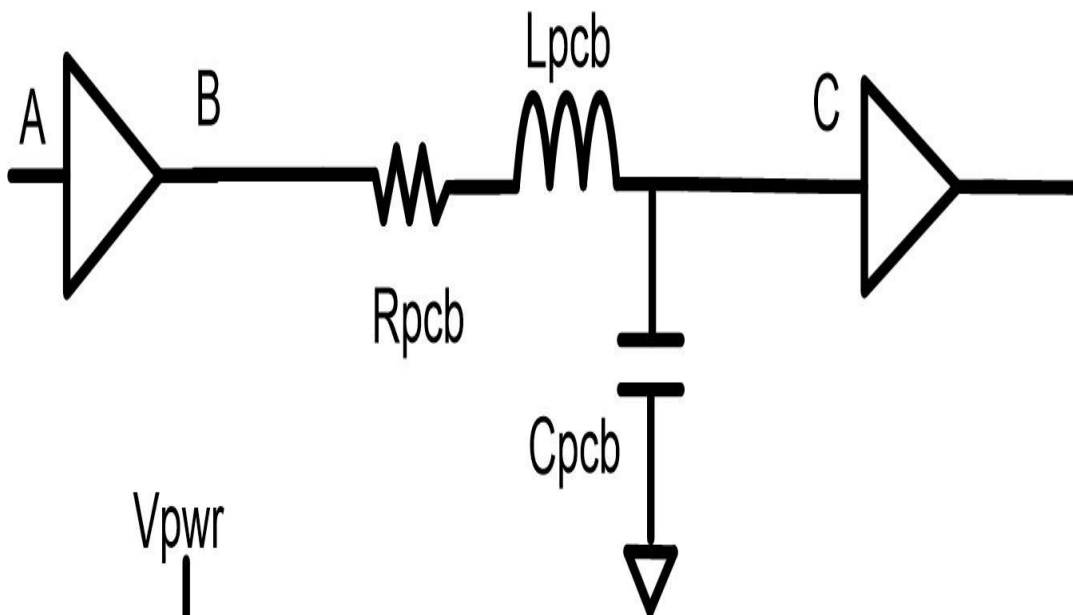


Figure 1-26. RC Loading in Digital Logic Interconnects

A simple driver and receiver pair are shown in **Figure 1-29-** . The middle diagram shows the interconnect impedance with resistance of the PCB (R_{pcb}), Inductance (L_{pcb}), and capacitance (C_{pcb}).

The bottom diagram includes an equivalent model for the output driver and receiver. For CMOS logic, the output driver can be modeled as a pair of switches and a resistor (R_{out}). R_{out} , depending on transistors used, will be typically 10 to 80 ohms. In addition, the receiver input behaves as a capacitance for most situations. This input capacitance is due to large, high current, ESD protection circuits, and 2pF to 10pF are common for each logic gate. Ignoring interconnect inductance, this model shows that an RC low pass filter is present in every digital interconnect.

With this more detailed model, the frequency limitations of logic are apparent.

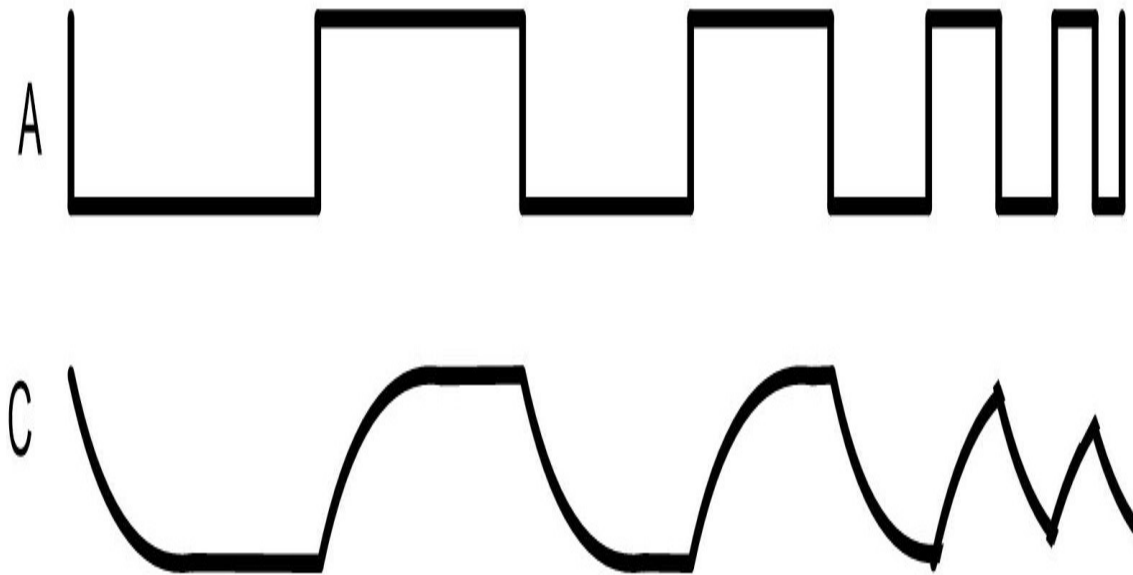


Figure 1-27. RC Loading and Signal Degeneration

Figure 1-29- looks at the limitations created by the RC circuits created by the output driver resistance and the attached capacitance. On the left, at

lower frequencies, enough time is available for the transient behavior to settle out before the next transition. On the right, as time is reduced between transitions, the device has insufficient time to settle out. The RC time constant of the system remains the same, but no well defined logic states are achieved. The system fails when V_{high} or V_{low} limits are violated, or setup and hold time on flip flops are not met. Ground referenced logic frequently fails when used for high frequency communication. Differential signals are commonly used at high data rates to minimize this limitation. More information can be found in the digital communication chapter.

Signal Integrity

Undesired effects can be created by a circuit in addition to its normal functionality. Many electronic systems create unwanted electronic noise by creating signals that radiate externally. Also, one part of a system can create noise that interferes with the proper functionality of a different part of the same system. In some cases, a poorly designed system can be sensitive to externally created electromagnetic interference. Signal integrity, Radio Frequency Interference (RFI), Electromagnetic Interference (EMI), Electromagnetic Compatibility (EMC), electronic noise, RF Immunity, electromagnetic shielding, radiated emissions, crosstalk, et al. are all related topics in this area.

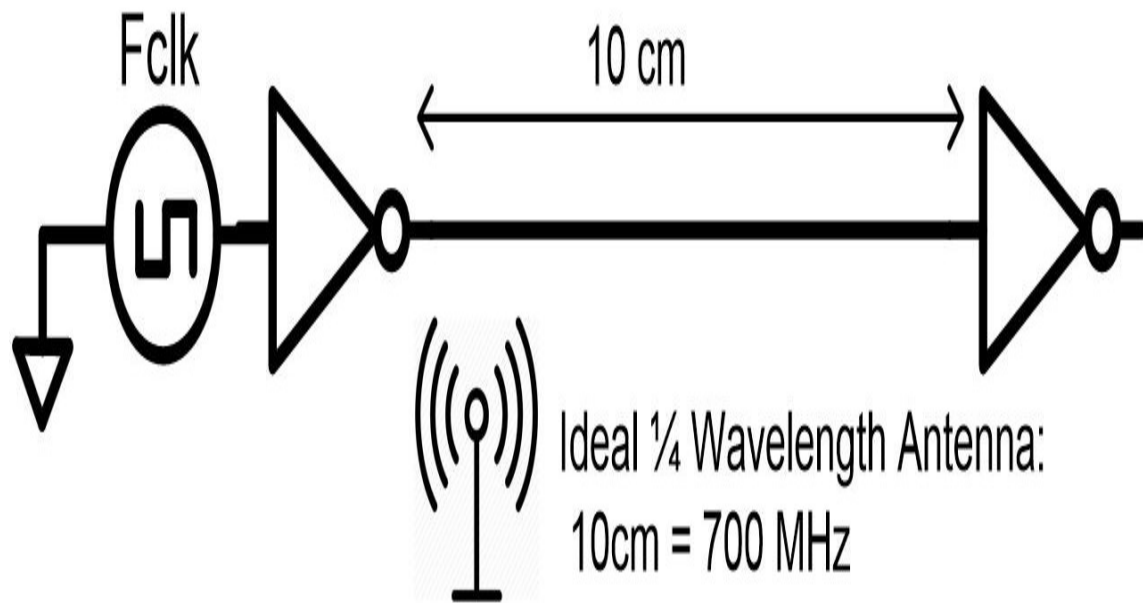


Figure 1-28. -x: Radiated Emissions

Again, using the 10cm connection discussed earlier, the above example creates radiated emissions with an ideal quarter wavelength antenna at about 700MHz. For this example, a 100 MHz clock would have its 7th harmonic have an optimal antenna, making the connection a radio transmitter. All frequencies would radiate from that connection, albeit with varying amounts of efficiency. Radiated emissions are especially problematic from computers and any device with large amounts of high frequency digital circuitry. The FCC (USA) and other worldwide regulatory bodies restrict the EMI that any device is allowed to produce in order to minimize interactive problems between electronic devices. More information can be found in the chapter on Noise Reduction Methods.

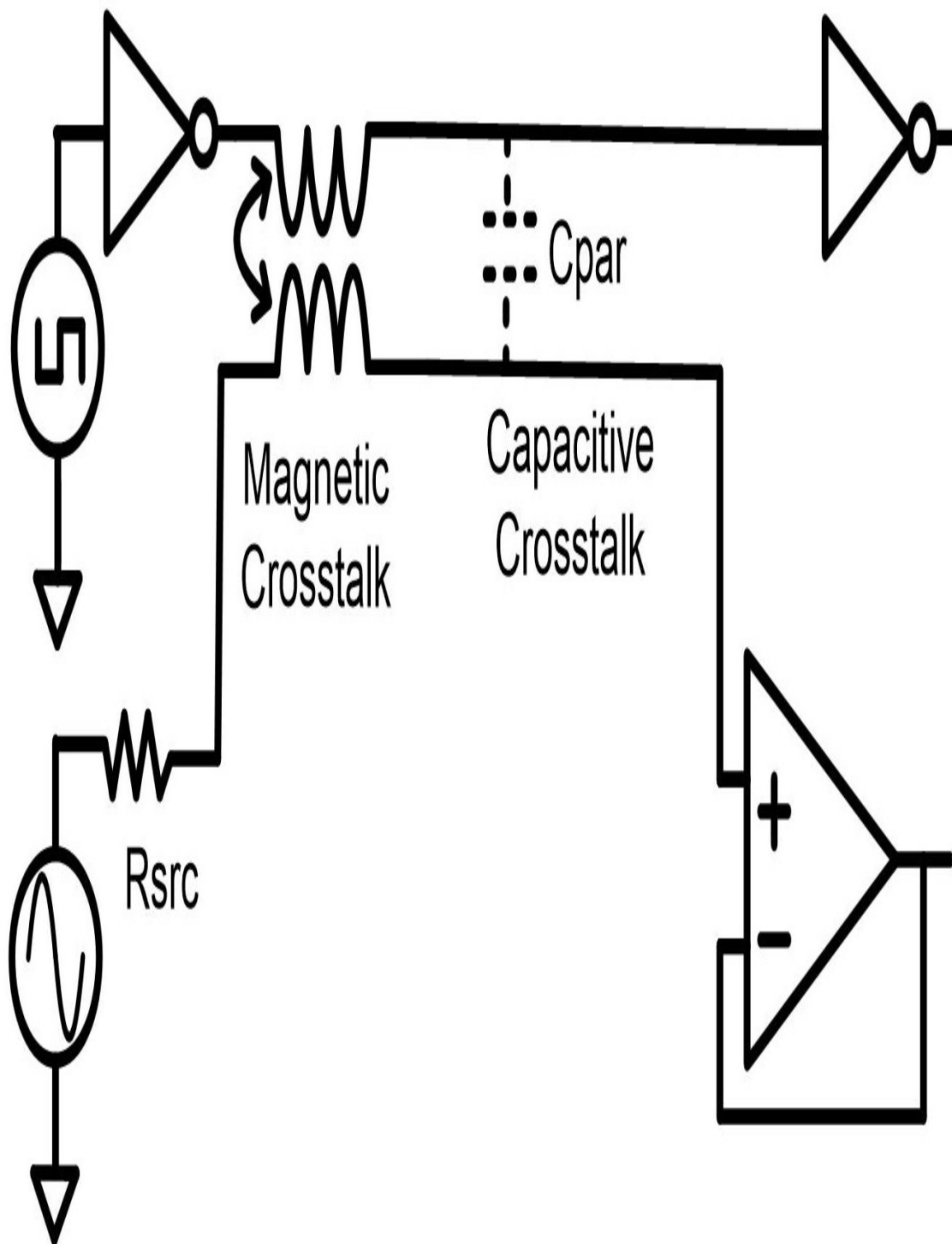


Figure 1-29. Sensitivity to External Signals

In addition to radiating EMI to the outside world, devices can be internally sensitive to unintended communication between circuits, commonly called

crosstalk. This noise coupling, can be capacitive or magnetic, and often takes both paths. **Figure 1-29-** illustrates a common problem with digital signals, having high frequency content, coupling to an unconnected circuit path. More information can be found in the chapters on digital communication, noise reduction, noise immunity, and PCB design.

Essential Concepts – Conclusions and summary

The important points of this chapter are:

- Academic simplifications often omit important details for the sake of clarity
- All connections have impedance
- Connection impedance can cause voltage variance across P&G
- Phase errors between receivers can exist even when driven by a common signal
- Capacitors includes ESL, ESR, self resonance, leakage, and voltage variance issues
- Resistors include ESL, EPC, and thermally generated noise
- High impedance resistors are susceptible to error due to environment issues
- Inductors include ESR, EPC, and self resonance issues
- Voltage and Current sources can have both impedance and noise components
- Switches have contact bounce, inductance, and variable contact resistance
- Op-Amps have limitations including: input/output voltage range, bandwidth, offset, output current, additive noise, and others

- Comparators have limitations in voltage range, input offset, response time and others.
- Power-ground instability can corrupt logic
- Asynchronous logic can create unpredictable behavior
- Ground based logic can be limited by RC loading that degenerates signal rise time and amplitude
- Digital devices often create radiated emissions
- Signals in one circuit can couple to another circuit without direct connection.
- Nothing is ideal

The intent of this chapter is to build awareness of the limitations and non ideal nature of electronics. Many engineers build awareness due to device failures from many of the items described here. Learning from experience is invaluable, but failure analysis is not the most efficient use of time or effort. Moving forward, the techniques presented in this book address these limitations while developing design solutions proven to work in real world applications.

Further Reading

- [Capacitor Guide—Dielectric Materials](#)
- [Ceramic Capacitor: C0G, X7R, Y5V, NP0, etc.—Electronics Notes](#)
- [Here's What Makes MLCC Dielectrics Different, Kemet Corporation](#)
- [Capacitor Types: Types of Capacitor—Electronics Notes](#)
- [High Reliability Capacitors—When the Mission Just Can't Fail, 2017, Wilmer Companioni, Kemet Corporation](#)

- Measurement of Voltage Noise in Chemical Batteries, Boggs, Doak, Walls, 1995, IEEE International Frequency Control Symposium
- Cascode Current Sources - Chapter 4, Analysis and Design of Analog Integrated Circuits, 5th ed. Gray, Hurst, Lewis, Meyer, John Wiley & Sons, ISBN: 978-0-470-24599-6

Chapter 2. Architecting the System

A NOTE FOR EARLY RELEASE READERS

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

This will be the 2nd chapter of the final book.

If you have comments about how we might improve the content and/or examples in this book, or if you notice missing material within this chapter, please reach out to the editor at arufino@oreilly.com.

Architecting the system defines all key parts of the design at the box level, and how to implement digital control at the core of the system.

A procedure to select a MCU is outlined. This chapter takes a big picture perspective and gets to the definition of all the needed boxes, while taking a more detailed look at criteria for the needed DCUs and their features. Later chapters will detail out the internal circuitry of those other boxes.

Options are explored in:

- what a “mostly digital” system is
- where digital controller units (DCU) are used
- where multiple DCU devices are needed
- differences between DCU, microcontroller (MCU) and microprocessor(MPU)

- alternative DCU methods: FPGA, CPLD, ASIC
- specialty methods for DSP and streaming data
- how to avoid data port bottlenecks

In this chapter, many topics touched upon are dealt with in more detail in later chapters. These items can affect the overall architecture so need to be briefly outlined up front to understand certain strategies. First, understanding some preliminary ideas and methods that need to be applied to all designs is important.

Preliminary Ideas

Defining a successful approach to a commercial product is the emphasis here. Certain common ideas need to be clarified at the start:

Simulate or Build

Generally there is no need to do a full simulation of most embedded controller systems. The controller code will be simulated within the IDE many times to get the control signals properly functioning. However, trying to simulate the functions created by those control signals probably isn't needed.

If the system is made up of SOC devices, simulating the circuits are generally not essential to create a first build. That work has been done by the IC designers of the SOC devices. Priorities and methods are very different for IC vs. PCB design. For an IC it makes sense to pursue accurate design simulations. IC fabrication is costly in both time and money. For a PCB, building out a first generation PCB is low cost, and quickly accomplished. This rapidly gets a real product in the hands of designers to debug, rather than trying to simulate the device. For a MCU with a simple set of peripherals another option is a development/demo board. These are available from most MCU manufacturers for minimal cost, and allow designers to go "hands on" quickly.

Obsolete – Through-Hole/Leaded Components

Circuit components with pins or wires that go through holes in the PCB are friendly to hand assembly and hobbyists. However, “through hole” devices, (also called “leaded” components) are not suitable for automated assembly or high frequency circuits. Also leaded components tend to be much larger than surface mount equivalents. Unless there is no other option, the use of leaded components is strongly discouraged. Any commercial design should be done using surface mount components for lower cost of assembly, better high frequency performance, and smaller physical implementation.

Obsolete - Discrete Gate Logic

The 7400 logic series, typically 4-6 logic gates per chip, (and many variants thereon), was started in the mid 1960's. It is still manufactured and widely sold. Herein, modern methods using programmed devices, and code driven logic functions are preferred. Singular digital devices should be seen only as level shifters and buffer-drivers.

Modern design strategies need to eliminate some other older and outdated practices as well.

Modern Design Strategies

As a general rule, analog signal processing should be avoided or minimized wherever possible. A typical example illustrates why this is a good idea:

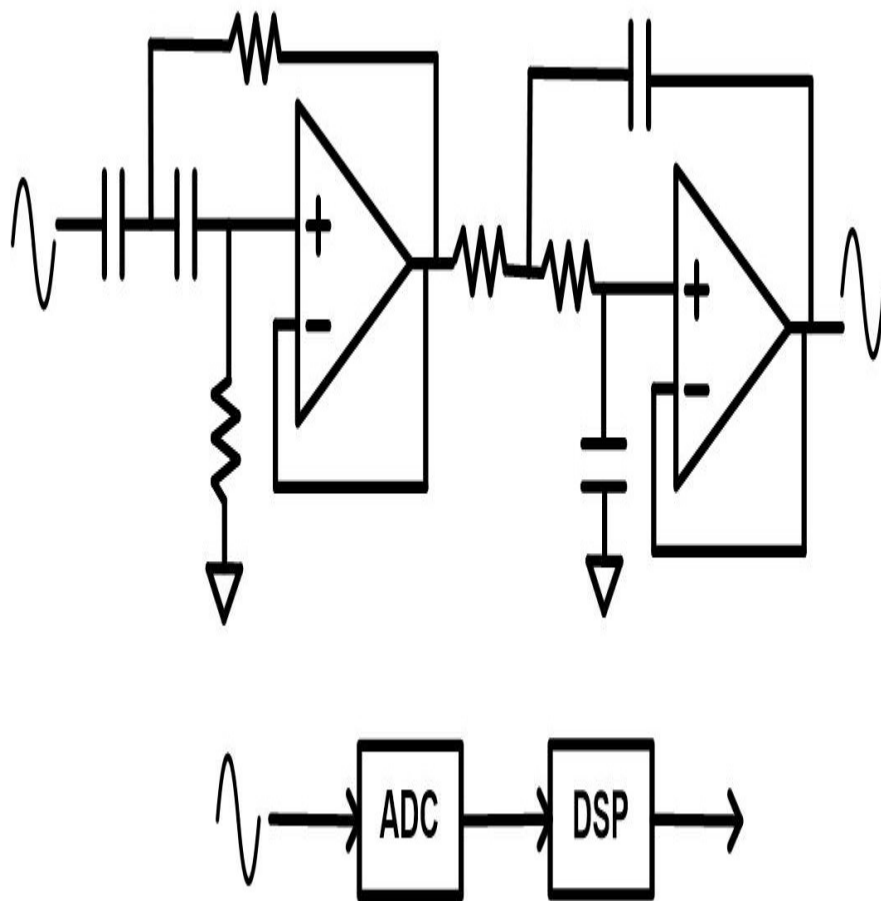


Figure 2-1. Analog Filters vs. Digital

Implementing a bandpass filter using op-amps can have many issues. The capacitors need to be high quality with good absolute accuracy and linearity. Accuracy of both the resistors and capacitors will affect frequency accuracy. Non ideal op-amps also contribute noise, distortion, and DC signal offset.

In this example, the op-amp bandpass filter requires 4 high accuracy plastic film capacitors, which are expensive. Using an ADC followed by a DSP (Digital Signal Processing) defined filter makes a low cost and very repeatable solution. Frequently an ADC built into the MCU and a SW defined filter will get it done. Independent ADC's are also available with a wide range of performance criteria and low unit cost.

There will always be analog signals to be dealt with however. The “analog world” is readily dealt with by a multitude of sensors. Getting those signals reliably processed is important.

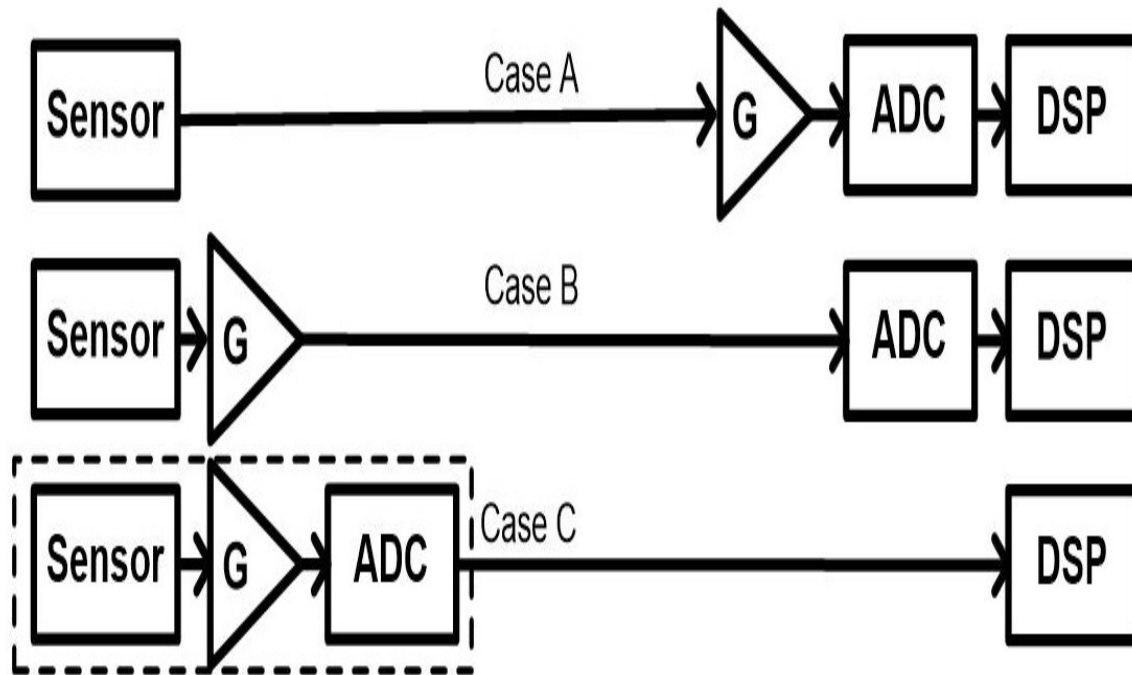


Figure 2-2. Quantize Signals Near the Source

As a general rule, transforming any analog signal into a digital equivalent quickly is a good strategy. Analog signals have issues with noise, distortion, and other signal integrity issues. Digital data streams are less problematic to connect over distance and can remain functional in a hostile EMI environment.

In the above example, Case A illustrates a sensor with a signal output that traverses a long interconnect before encountering a gain amplifier. The signal amplitude output of the sensor, depending on the device, could be anything from microvolts to volts. With small amplitude signals, the sensitivity to external EMI can be high. Placing an amplifier in close proximity to the sensor, Case B, improves this situation. Ideally, to eliminate long path analog signals, Case C moves the ADC near to the

sensor as well. More and more sensor devices are being developed where the sensor, amplifier and ADC are implemented in a single package with a serial port digital interface.

Control systems are another area where analog methods are outdated.

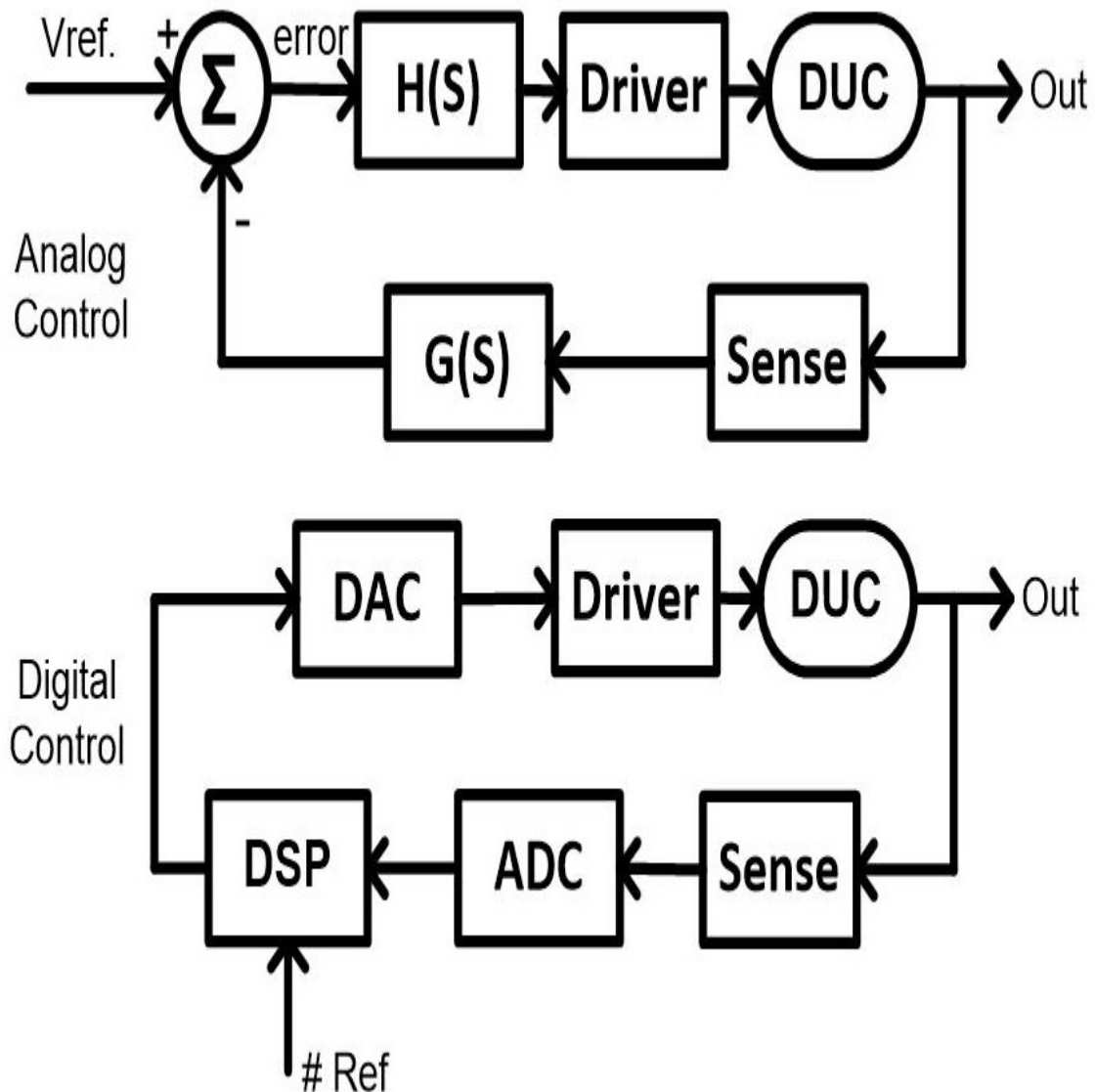


Figure 2-3. Control Systems Classic Analog vs. Digital Methods

A classic analog control feedback loop is shown above. A Devices Under Control (DUC) is controlled through a device Driver and a linear system of forward transfer function, $H(S)$, a sense system, feedback transfer function,

G(S), and an error differencing amplifier. Although this type of control feedback system exists internally within many analog and mixed signal IC's it is rarely used anymore at the system level in modern designs.

Modern control feedback systems are mixed signal in nature, with some analog characteristics within the Driver, DUC, and Sense blocks. However, the DSP functionality can take many forms to achieve optimum control.

Also, communication and control between PCB's in a multi board system needs to be structured for EMI immunity.

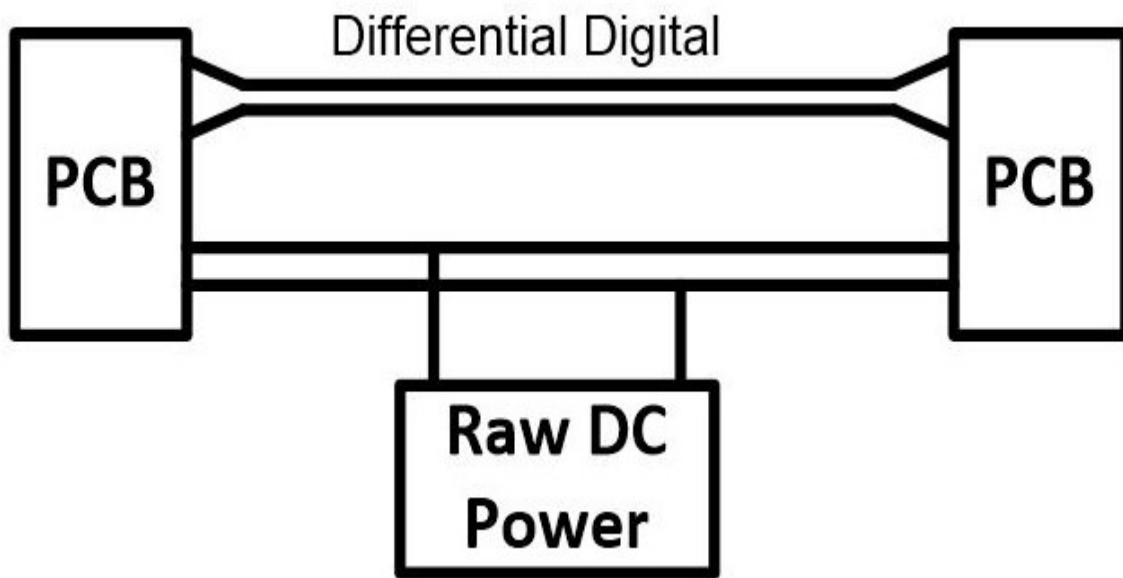


Figure 2-4. Communication and Power Between PCBs

Controls and signals between PCBs, consists of sharing ground, power, and data. The data communication should use a differential digital system with error checking capability. This approach avoids most noise problems and signal integrity issues. Also, this method tolerates variation between the grounds of each PCB. Power on each PCB is locally regulated allowing a large amount of variance in the “Raw DC Power” before on board functionality is degraded. Remember that all connections between PCBs have impedance and will cause voltages to vary.

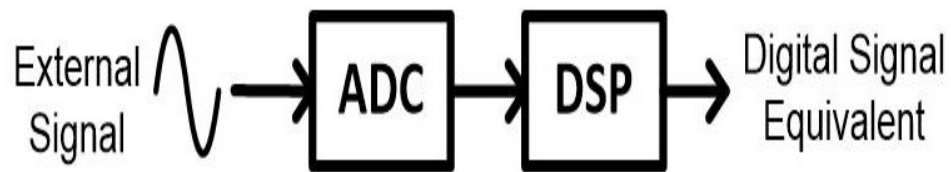
Understanding what a mostly digital approach looks like and what are the limitations to the method is next.

Mostly Digital Design

The mostly digital system is a simple concept:

Convert any analog signal (control voltage, sounds, video information, etc.) to a digital equivalent data stream. Process that data stream using digital methods only. Convert the data stream to an analog format only if needed as an output.

Industry frequently refers to this as a “Mixed Signal” approach.



Possible DSP Functions:

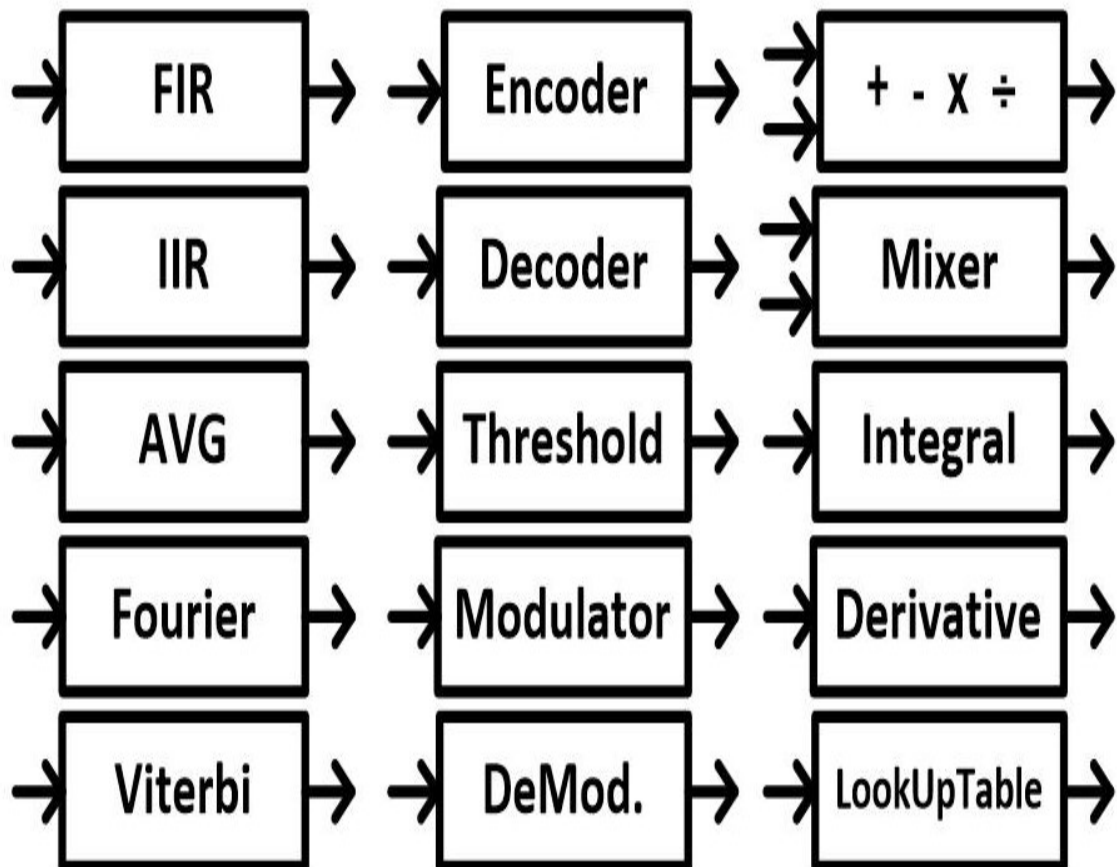


Figure 2-5. Common Structures Used in a DSP Strategy

Any form of signal processing that can be defined as a math function or some type of input-output relationship, can be created in a DSP system. Some examples:

- Mixer (Multiply), Divide, Sum, Difference, Integral, Derivative
- Non Linear Transfer
- Closed Loop Control - Proportional-Integral-Derivative (PID)
- Closed Loop Control - Fitted Trajectory
- Encoder or Decoder
- Amplitude Qualification, Threshold or Comparator
- Modulator or Demodulator
- Filters FIR or IIR (LPF, HPF, BPF, et. al.)
- Averaging and Smoothing (AVG)
- Time Domain to Frequency Domain (Fourier) Transform
- Viterbi Detection

If you can accurately quantize a signal the possibilities of DSP are endless. Consequently, ADC performance can limit this method in some cases.

DSP Methods - Versatility & Limits

In order to get better signal to noise performance, a higher number of ADC bits are needed. In order to accurately sample a signal you need to convert at a rate faster than twice the signal bandwidth.

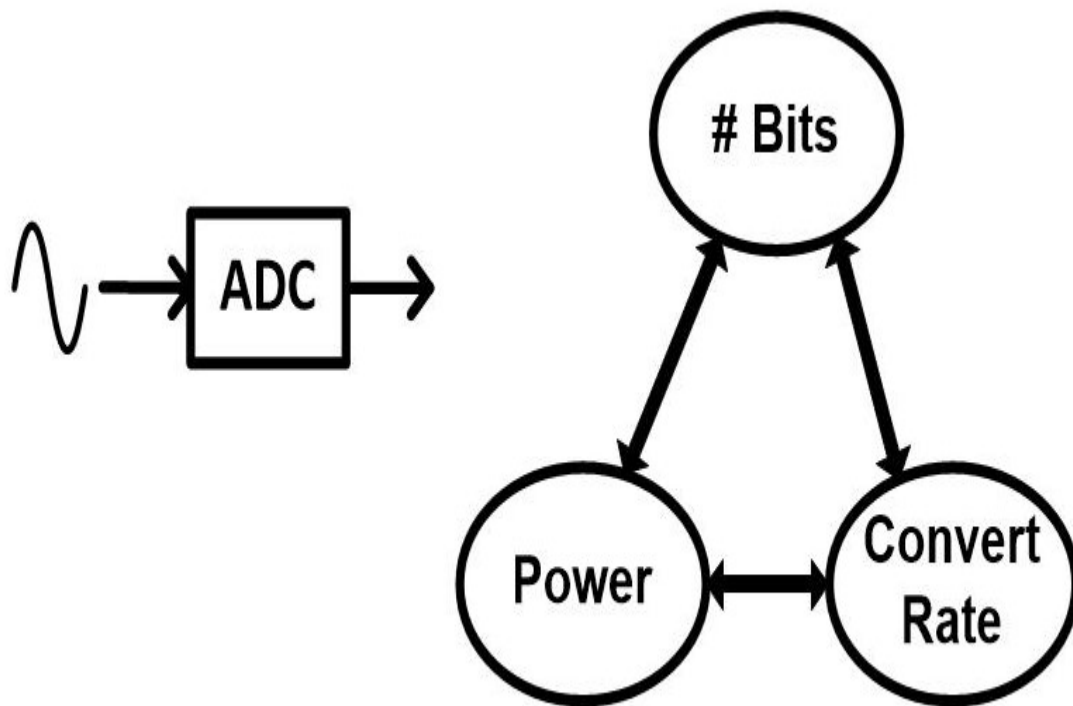


Figure 2-6. Resolution and Conversion Rate Limits of ADC

The internal circuitry of an ADC has performance limitations. Power consumption, the number of bits, and the conversion rate interact in a design. Increasing the conversion rate or increasing the number of bits, makes the power consumption go up. Designing a “fast” converter, with a “large” bit count and “low” power consumption can be challenging, but a steady improvement on “fast – large – low” performance is still ongoing. What was difficult to achieve a few years back is now readily available.

As of 2022, ADC’s are available with resolution from 4 to 32 bits, and sampling rates from 4/second to 6.4G/second. Many of the very fastest devices are expensive, power hungry, and not suitable to consumer applications. However, many cost effective options are available under 500M samples/sec. More information can be found in the chapter on sensing peripheral devices.

A simplified case study of a radio system illustrates why in some situations analog methods are still needed.

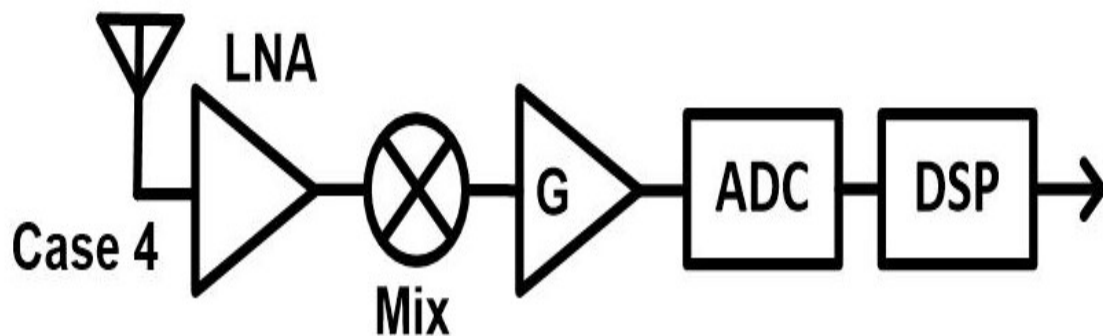
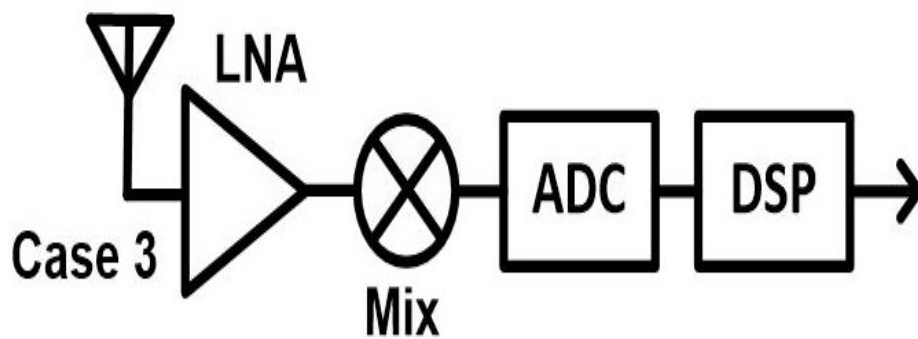
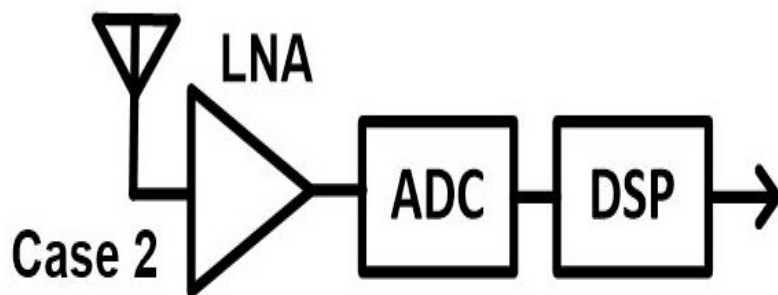
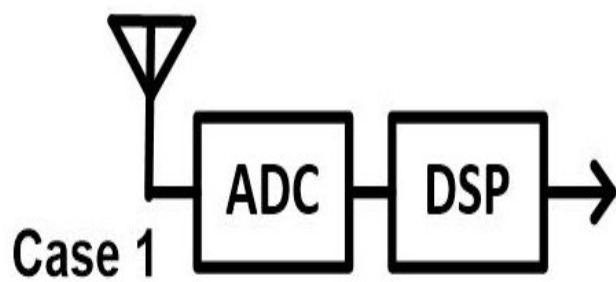


Figure 2-7. Software Defined Radio, Antenna Connected to ADC

The “Software Defined Radio” is a simple idea: Connect an ADC to an antenna and perform all the signal processing in a DSP block. Case 1 illustrates that idea.

On paper, Case 1 looks easy. The specifics of implementation can be a different story. The dynamic range of a typical radio receiver is 90 dB. To support that dynamic range would require 16 bits, or more, resolution. State of the art, 2022, ADCs with 16 bits resolution top out at a maximum of 500M samples/sec, so this approach may be usable up to a 250 MHz carrier frequency. Most cell phones operate from 900MHz to 2GHz, so it’s not going to work for cellular.

Case 2 introduces a Low Noise Amplifier (LNA) between the antenna and ADC. Adding 20dB of gain drops the needed resolution down to 12 bits. Again looking at devices available in 2020, there are options available to 6G samples/sec but these are expensive and very power hungry, more suitable to rack mounted instruments than handheld cellular. But, it could be done for RF signals up to 3GHz, with a power cord connected and a high price to implement.

Case 3 and Case 4 introduces a mixer, converting the RF signal down to a low frequency baseband. The high frequency of the RF carrier is no longer an issue so only the bandwidth of the signal modulation needs to be converted. For these situations, the signal of interest is under 100MHz, and frequently much less. At lower conversion rates a multitude of suitable converters are available with battery friendly power levels and suitable bit resolutions. Most cell phones now use methods similar to this. This is a simplified example, but illustrates the need for analog methods in some situations.

Audio systems are friendly to the mostly digital approach.

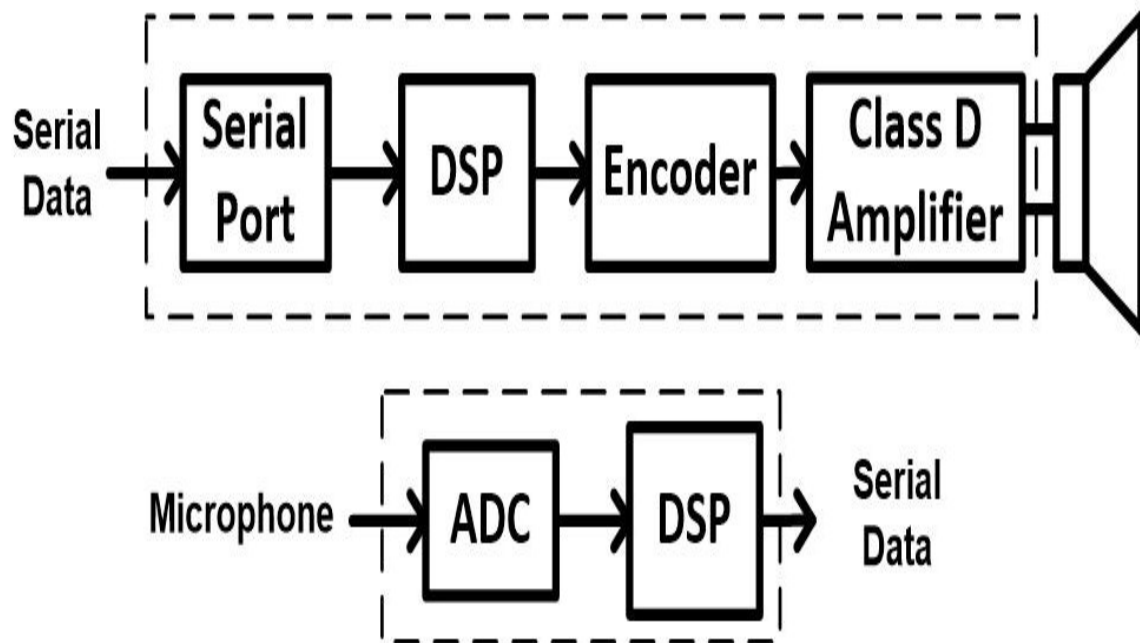


Figure 2-8. Mixed Signal Audio

Modern audio, on the inside, is largely digital. Many microphones are fed directly to an ADC and turned into a serial data stream. High quality recording microphones are now available with digital USB outputs that connect directly to a computer.

Recording studios use analog differential signal microphones and XLR connectors largely due to legacy infrastructure. Modern mixer boards immediately ADC the microphone input and digitally process the audio data stream. As shown in the illustration, audio output can be kept digital up to the speaker's power amplifier that uses a digital input. The encoder driving the amplifier converts the binary data to a pattern suitable for driving a switched amplifier.

Sensors come in many forms to interface to the analog world.

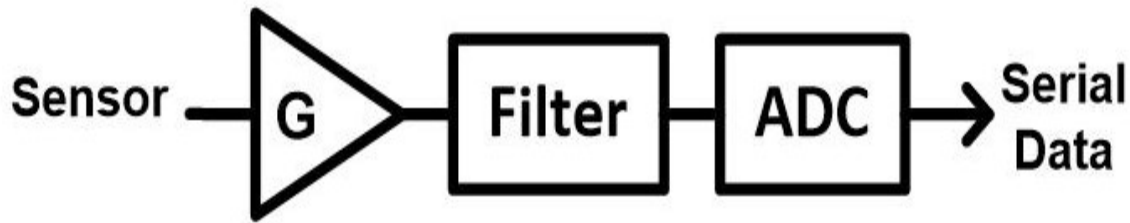


Figure 2-9. Analog Signal Processing, Signals and Sensors

A sensor output may require amplification to use the full range of the ADC. Also, the characteristics of the sensor output may need to be filtered by analog methods before connecting to the ADC in order to avoid aliasing. More specifics in the chapter on sensing peripherals.

Digital Control Methods– DCU, MCU, MPU

An “Embedded System” or an “Embedded Controller System” can be loosely defined as an electronic or electromechanical system which uses a digital controller to control and sense things in the system. The digital control unit (DCU) is a generic concept. The DCU can be a microcontroller (MCU), a microprocessor (MPU), or a more specialized digital methodology. Digital control can be either hardware based, as a Boolean state machine or software based, as in code written for a computer. Compiled code that was written for and downloaded into an MCU is known as “firmware”.

An embedded system is single purpose, unlike a multipurpose computer that has the capability to change programming. An example is useful:

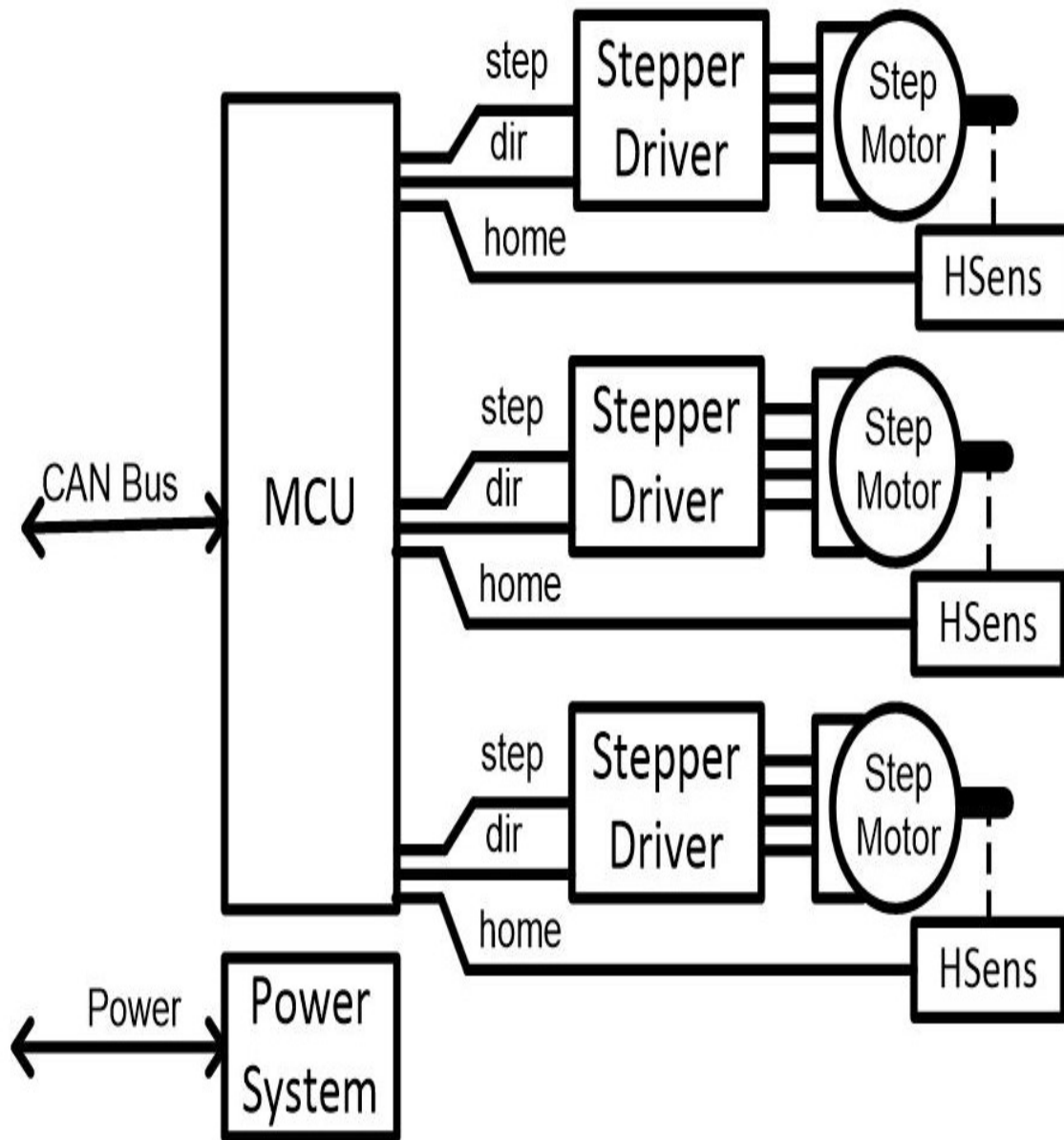


Figure 2-10. Embedded Controller System Typical Example

The above shows a motor driver-controller PCB using an MCU to control three stepper motors. The MCU issues step and direction commands to the motors, and has a sensor for the home position. Also, the MCU communicates to a remote master over a CAN Bus, to receive commands and send status. This structure is used in a robot for motion control where dedicated MCUs control groups of motors.

Terminology in MCU & MPU Specifications

As of 2023 there are about 40 semiconductor manufacturers of microcontroller chips. There are thousands of MCU and MPU variants available. While sorting through specifications, the specialized jargon can be confusing. A brief terminology primer is a good start.

Harvard Architecture

A CPU that uses separate data-address buses for instruction memory and data memory. Most MCU devices are Harvard Architecture.

Princeton Architecture (aka von Neuman Architecture)

A Princeton architecture CPU uses common address-data buses for both instruction memory and data memory.

RISC (Reduced Instruction Set Computer)

Typical RISC CPUs limit the instruction set to reduce the complexity of the physical design. This also keeps execution time of an instruction within a single clock cycle. Pipelining of instructions and resulting data is often used to do this. Most MCU devices are RISC devices.

CISC (Complex Instruction Set Computer)

A CISC processor executes more complex instructions than a RISC device. The compiler for a RISC device breaks that complex instruction into simpler instructions suitable to the RISC instruction set. The HW of a CISC machine needs to be more complex. Most present day processors use RISC architecture, except for Intel's X86 devices.

X86 Processor

The X86 architecture includes the Intel 8086, 80186, 80286, 80386, 80486, and others. The X86 series is a CISC architecture that has been the foundation of Intel-Microsoft personal computers. Generally, X86 devices are not used in MCUs within embedded systems.

ARM (Advanced RISC Machine, or Acorn RISC Machine)

A series of RISC processors, with 32 bit data and 26 or 32 bit addressing. ARM processors are licensed as intellectual property, by ARM Holdings, to many semiconductor manufacturers, and are widely used. Chip manufacturers add peripheral support devices creating a finished MCU.

MIPS (Microprocessor without Interlocked Pipeline Stages)

A RISC processor that uses a pipelined architecture. MIPS devices are optimized for streaming of data, and are available from multiple vendors as a standalone MPU, or used as an intellectual property design by semiconductor manufacturers. MIPS cores are popular in general data processing and applications that process data streams.

8051 and 6502

The 8051 is an early generation, 8 bit data, 16 bit address, MCU from the early 1980's. The device is still used in many simple controller applications, implemented as a digital equivalent on a more modern CMOS process. A multitude of variants with different features exist, from many vendors.

The 6502 was developed around 1975 and is similar to the Motorola 6800 MPU of that era. The 6502 uses 8 bit data and 16 bit addressing. The 6502 is also still available in more modern CMOS versions.

Both the 8051 and 6502 are programmed at a low level, using Assembly language programming.

PIC, AVR, ATmega

These are several different families of microcontrollers, that started as 8 bit Harvard architecture, RISC processors. Some of these have expanded their product lines to include higher performance devices.

Hardware (HW) Controllers

Hardware based controllers include the PAL-PLD, CPLD, FPGA, Digital ASIC (metal layers) and Digital ASIC (all layers). A brief look at each of these:

PAL (Programmable Array Logic) and PLD (Programmable Logic Device)

These are both programmable logic arrays. Devices defined as a “PAL” tend to be 28 pins or less, with 5V compatible interfaces and most have through-hole mounting. PLD devices are smaller variants of the CPLD, although no industry standard naming conventions exist, so this can vary. PAL use is being phased out, and is discouraged for new designs.

CPLD (Complex Programmable Logic Device)

The CPLD ranges in size from 20 I/O ports to over 200. Internally the CPLD uses a programmable interconnect on a field of gates or macrocells. If a hardware programmable approach to a DCU is needed, a CPLD is often capable of getting it done.

CPLD and FPGA devices are generally configured using Verilog or VHDL, to define functionality.

FPGA (Field Programmable Gate Array)

The FPGA ranges in size from 10 I/O ports to over 1900. The FPGA uses programmable lookup tables to form logic functions. Complex, large array, dedicated logic can be done within an FPGA.

Frequently, an FPGA is used in a first generation product, and a digital ASIC is then created when the design has been proven.

Digital ASIC (Digital Application Specific Integrated Circuit)

The digital ASIC can't be programmed, so is not suitable for initial product development. A functional digital design needs to be worked out using another platform before fabricating an ASIC. If the product is high volume and needs fast logic, this is the preferred method.

Two variants of the digital ASIC exist. The first uses unique metal layers on a standardized array of fabricated transistors. The second is the all layer implementation where transistors and metal interconnects are all unique to the device. Both are widely used.

HW based controllers are suitable when product volumes are large, or fast dedicated logic is needed. The majority of controllers are readily implemented using a software based controller.

Software (SW) Controllers

Software/Firmware based control includes micro controller (MCU) and microprocessor (MPU) devices. Differences between MPU and MCU features can be fuzzy. A MCU tends to have more features to make it self sufficient over a MPU. But there can be a lot of feature overlap.

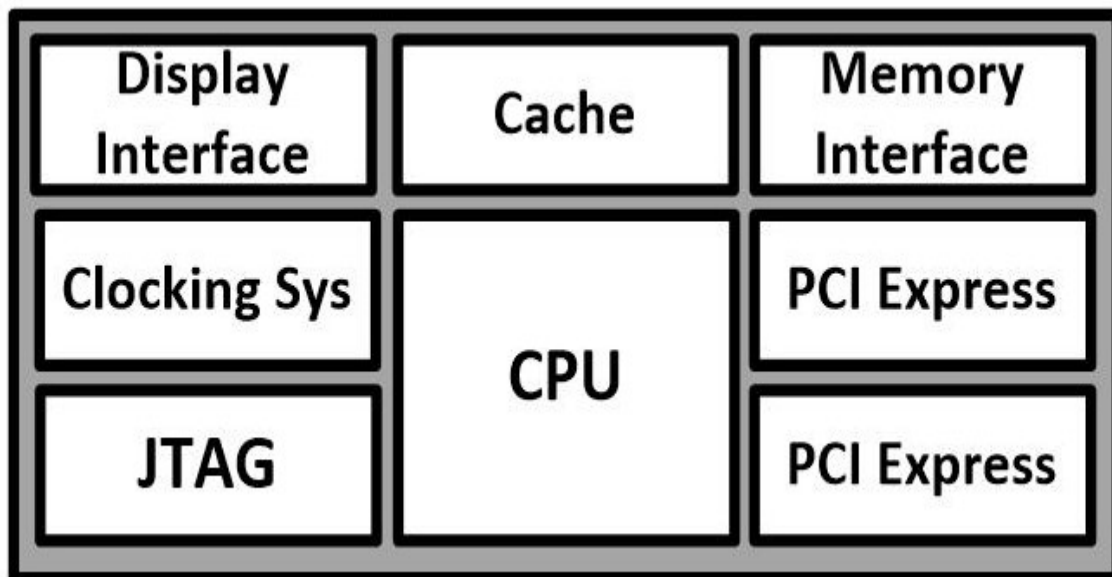


Figure 2-11. Typical Microprocessor (MPU) Feature Set

As shown, most MPUs keep the RAM and ROM outside the chip with an emphasis on interfacing to external memory and peripherals. Many state of the art MPUs have advanced to a greater number of internal features than shown here, but this illustrates the concept.

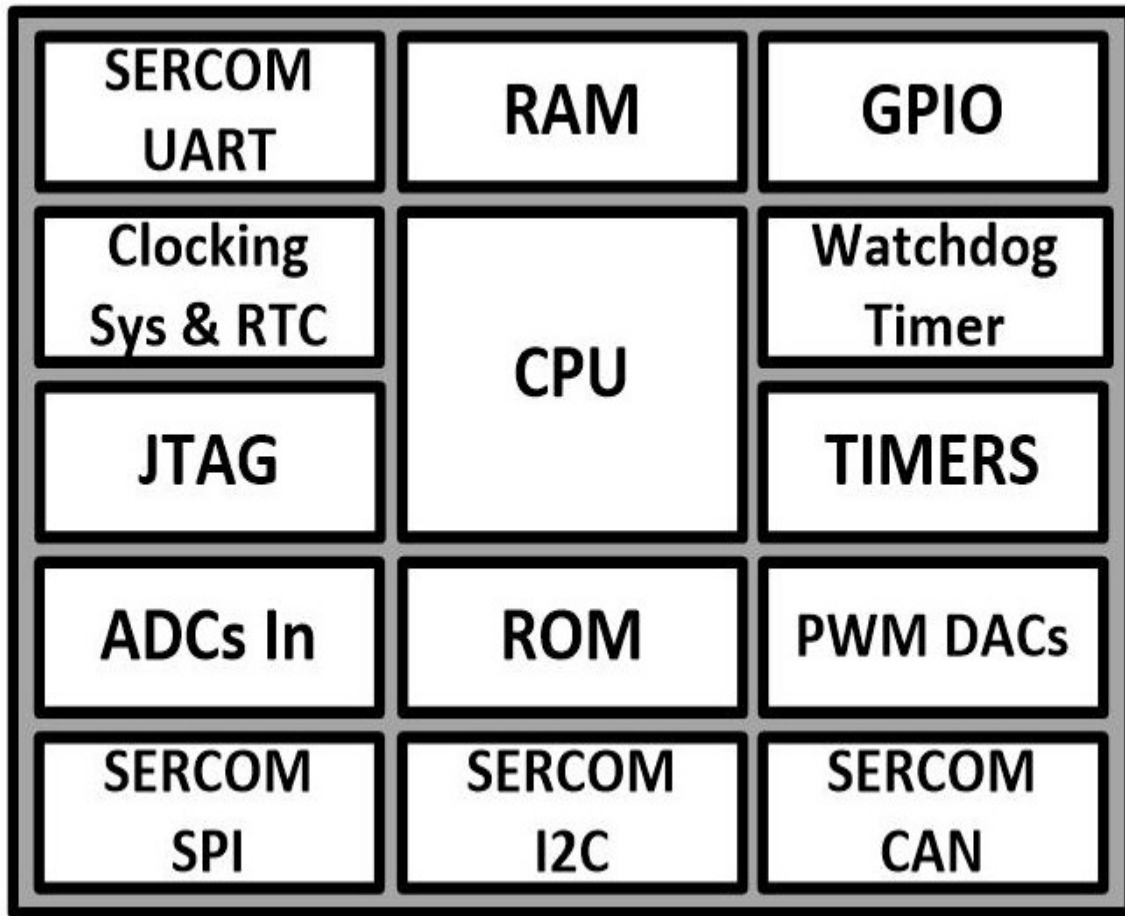


Figure 2-12. Typical Microcontroller (MCU) Feature Set

Inside the typical MCU are the features needed to allow the device to function with minimal external support. The working memory RAM and the control program memory ROM have been internalized. Serial communication to external peripherals is supported by UART, SPI, I2C, and CAN interface capability. Both ADC and DAC capability are common features to allow digital-analog interfaces.

Frequently a clocking system includes clock multipliers and a RTC (Real Time Clock) oscillator. Internal timers are commonly used to time or control external functions, and a watchdog timer is often provided as a safety reset feature. GPIO, (General Purpose Input Output) is also available to control and monitor individual control lines.

There are thousands of different MCU and MPU devices on the market. These are two simple examples for illustration, whereas there are many

possibilities to pick from, with features ranging from simple to multi core devices with a diverse and extensive set of options. This chapter provides the background knowledge to navigate the options maze.

Computers vs. Controllers

A multitude of single board computers are readily available, and many vendors also produce development boards for their MCU devices. Off the shelf boards are one way to start, but may not be a good final fit to what the project needs.

Single board computers are configured to drive typical computer peripherals and a MCU board is typical set up to interface to other ICs and provide stimulus/response through individual ports. A comparison is useful.

Raspberry Pi (MPU) vs. Arduino (MCU)

Two very popular single board devices illustrate typical differences between an MPU and MCU. The Raspberry Pi can be considered a single board computer and the Arduino is a microcontroller development board.

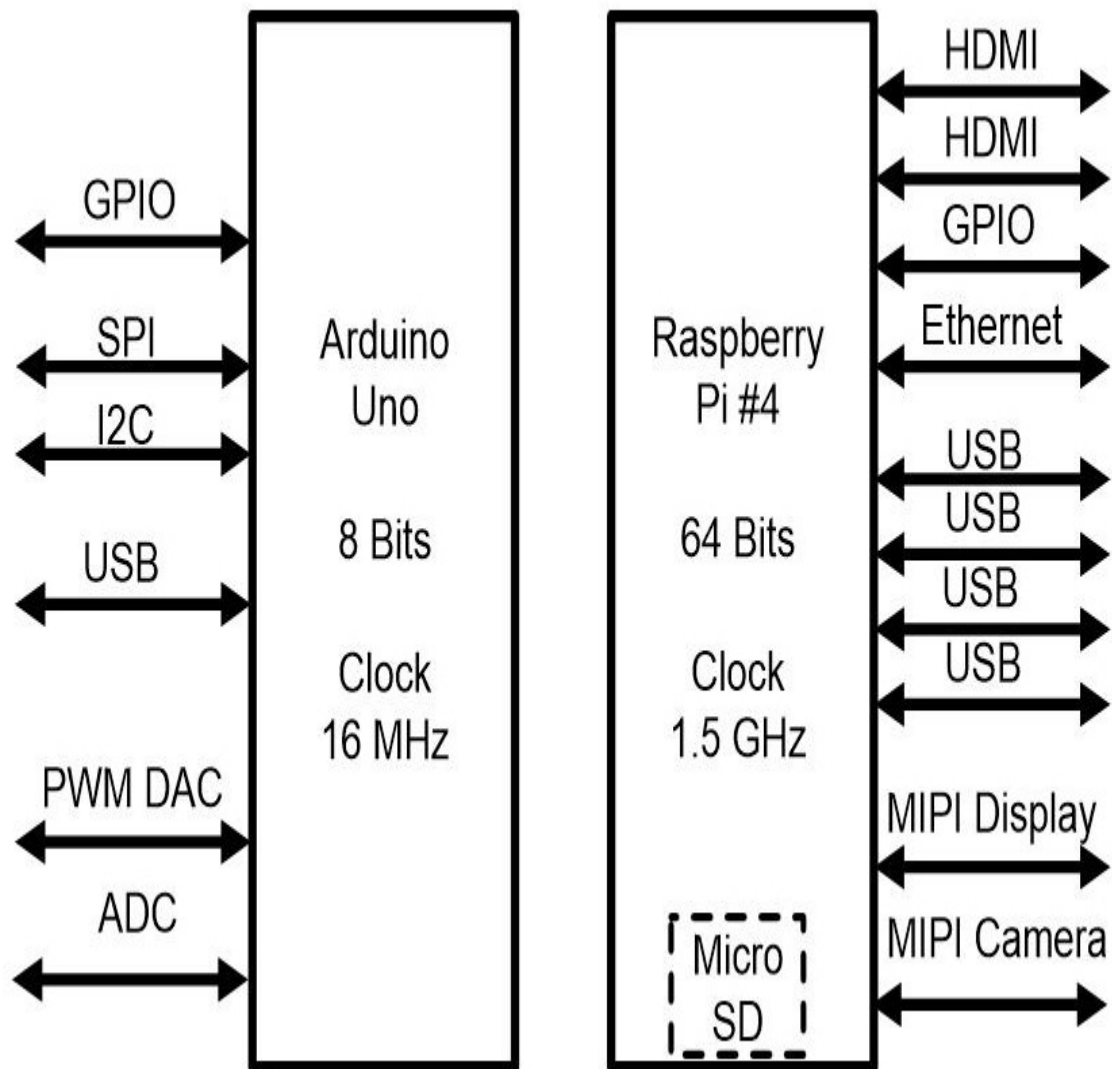


Figure 2-13. Arduino MCU vs. Raspberry Pi MPU

The processor specifics, and what is on the board, tells the story. The MCU runs on a 16 MHz clock, and the MPU uses a 1.5GHz clock. The MPU is set up with video display, USB, and Ethernet interfaces. The MCU is configured with interfaces for peripheral ICs using SPI and I2C interfaces. Analog friendly ADC and DAC interfaces are available on the MCU. The MCU has an internal 8 bit structure, and the MPU has 64 bits.

The MCU is optimized to be fast enough and have enough processing power to control the slow functions of an electro-mechanical system. The MPU is configured for speed and enough processing power to run an

operating system, feed graphics to two video ports and interface with multiple peripherals over multiple I/O interfaces. The differences in the target applications are clear.

Multi Purpose and Specialty MCUs

Frequently, MCU cores are bundled with a specialized block of logic.

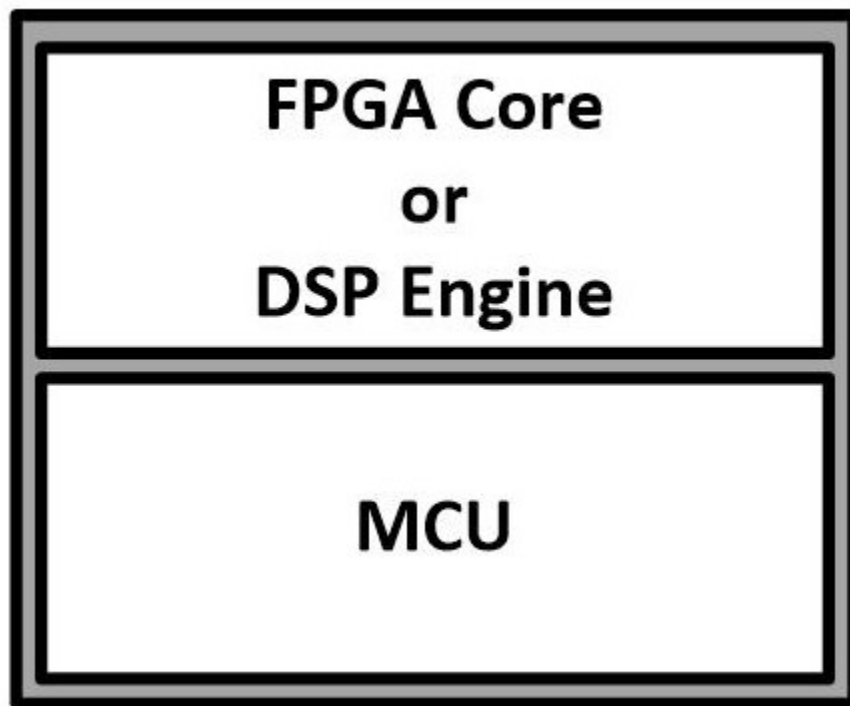


Figure 2-14. DSP or FPGA Core Combined with MCU

Two of the most commonly bundled implementations are an FPGA gate array or a dedicated DSP engine. The FPGA/DSP core can run at a high speed, processing a data stream, while the MCU serves as a control and monitor to the process. DSP cores are optimized for the floating point math of DSP algorithms.

A multitude of other devices are being bundled with an MCU for specialty applications.

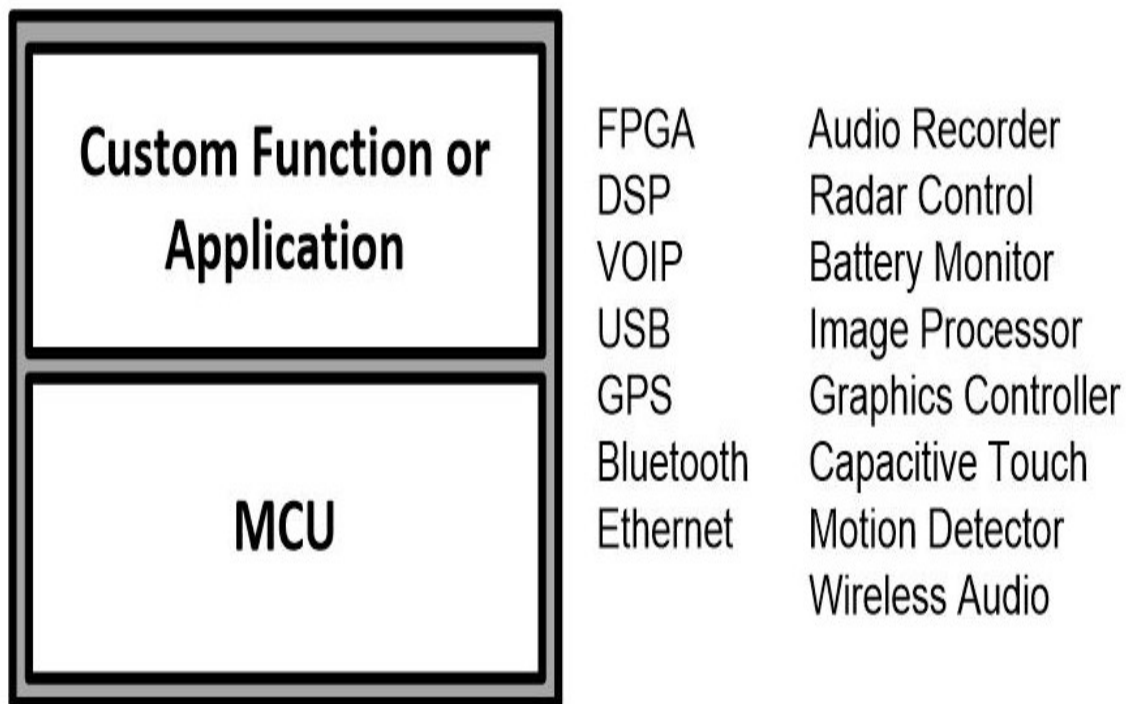


Figure 2-15. Examples of Specialty Devices

The above is a small sample of specialty devices that have been bundled with a MCU. If sufficient demand exists, some manufacturer has probably produced a specialty device to fill the market. Single chip solutions are popular with consumer products where volumes are high and low cost is important.

Highly complex devices have also been created using chip sets for select products.

Chip Set Methods

A chip set is a group of ICs designed to function together in order to create the electronics for a particular product. A chip set needs minimal external support with the devices designed to be plug together compatible. Chip sets have been used since the early 1970's. The original Intel 4004 MPU was part of a 4 chip set for the electronics of a calculator. A chip set is often

developed when a high volume demand exists, or small form factor is mandatory.

One of the most widely adopted chip sets is the one used on the personal computer (PC) motherboard.

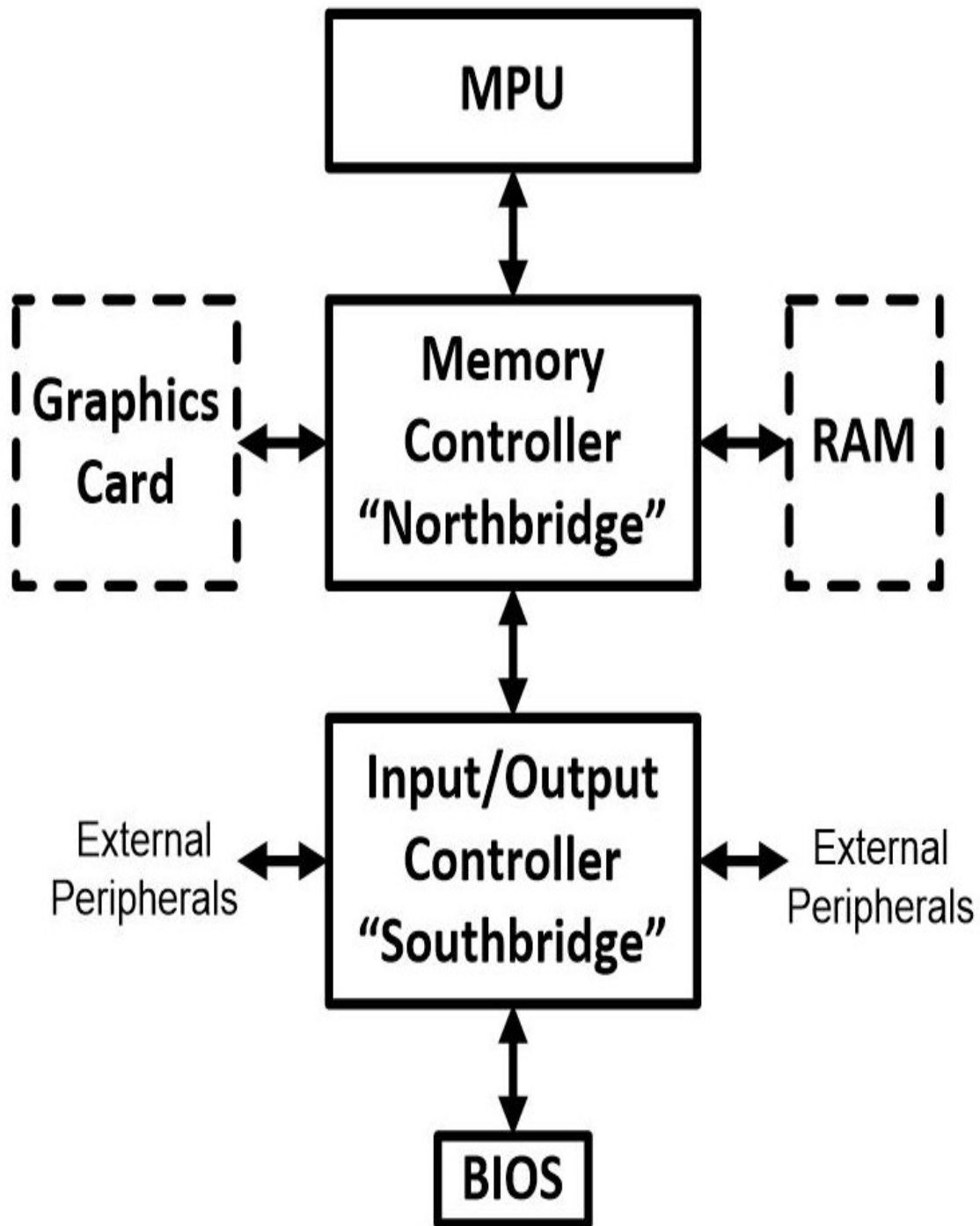


Figure 2-16. Computer Mother Board Chip Set

Many PCs used a four chip system of: MPU, memory controller (Northbridge chip), I/O and peripheral controller (Southbridge chip), and BIOS. This was the center of a PC motherboard. Depending on the

particular architecture, some newer designs have reduced the chip count further.

The smart phone requires a very high level of integration to achieve both low power and compact size.

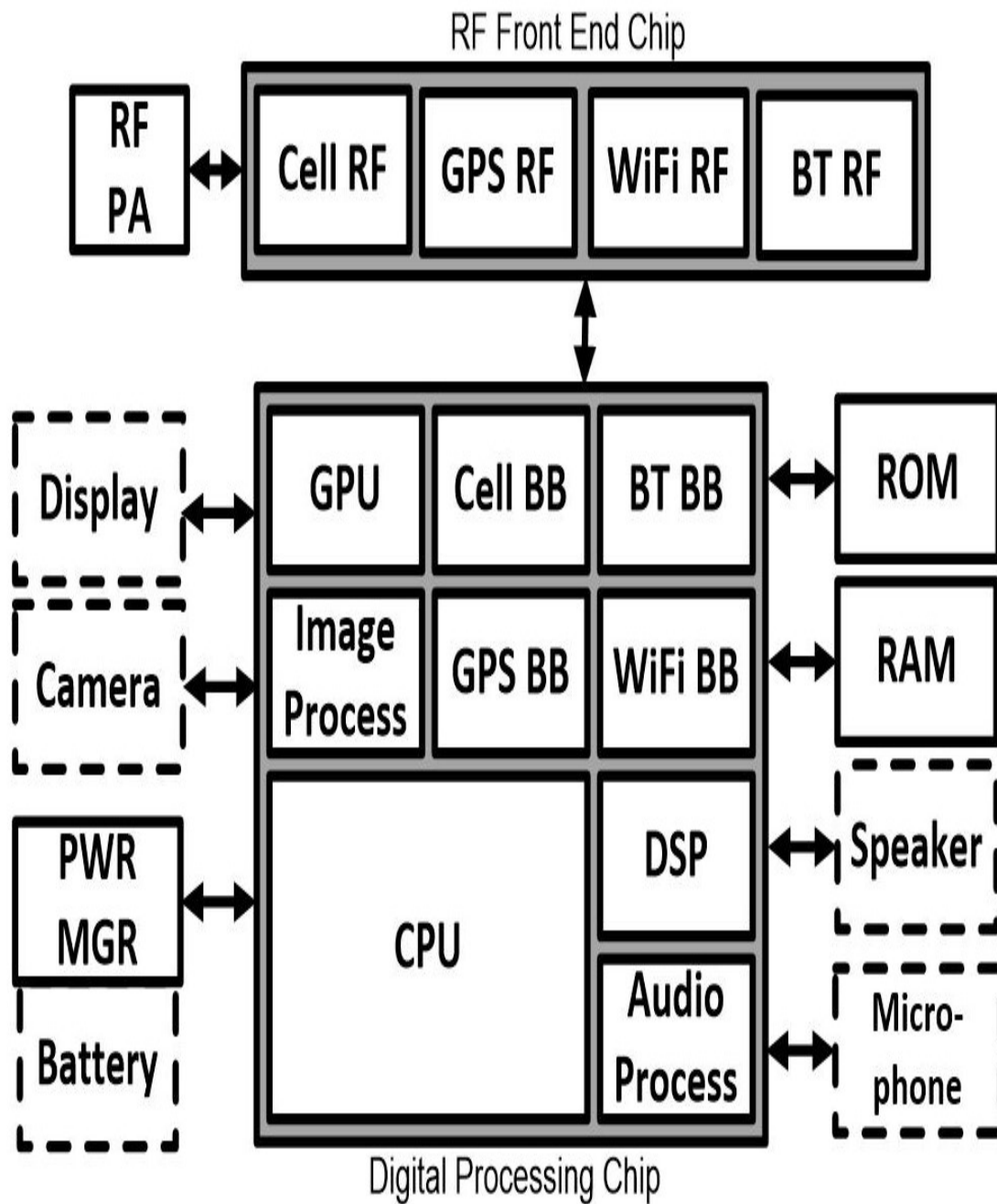


Figure 2-17. Cell Phone Chip Set

Most smart phones use a chip set similar to the above. Most digital signal processing and control exists on a single digital ASIC. This ASIC uses a silicon process optimized to high density logic. Supporting that are chips to provide other needed functions. These are often fabricated using a semiconductor process better suited to their application. The PWR MGR (power management chip) supports battery management and charging. The RF Front End chip is fabricated on an “RF friendly” semiconductor process. The RF Power Amplifier (RF PA) is often done using a specialty process (Gallium Nitride or Gallium Arsenide) suitable to the needed RF output power.

Other products that embrace the chip set approach includes digital cameras, HDTV systems, and set top boxes.

System Architecture Options

MCU and MPU devices see a countless number of applications. Not all are single processor, and frequently it's not evident that there's a MCU in the device. Taking a look at some common applications gives insight into some typical MCU use scenarios:

A MCU with some additional special feature content is frequently implemented as a stand-alone device:

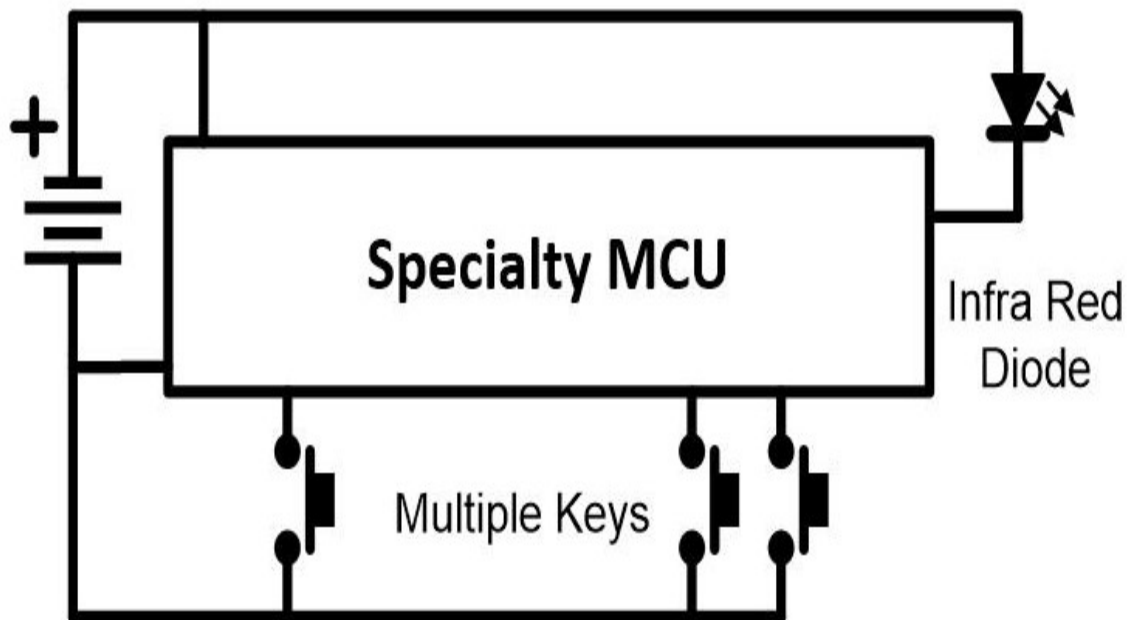


Figure 2-18. All in One SOC Device

This example has pushbutton switch inputs, a battery, and an infra-red LED. This example exists in TV remote controls. Similar but different, another implementation includes an internal RF transmitter, and an antenna instead of the LED, and exists in many wireless key devices. The above uses specialized silicon bundled with a generic MCU to create a solution.

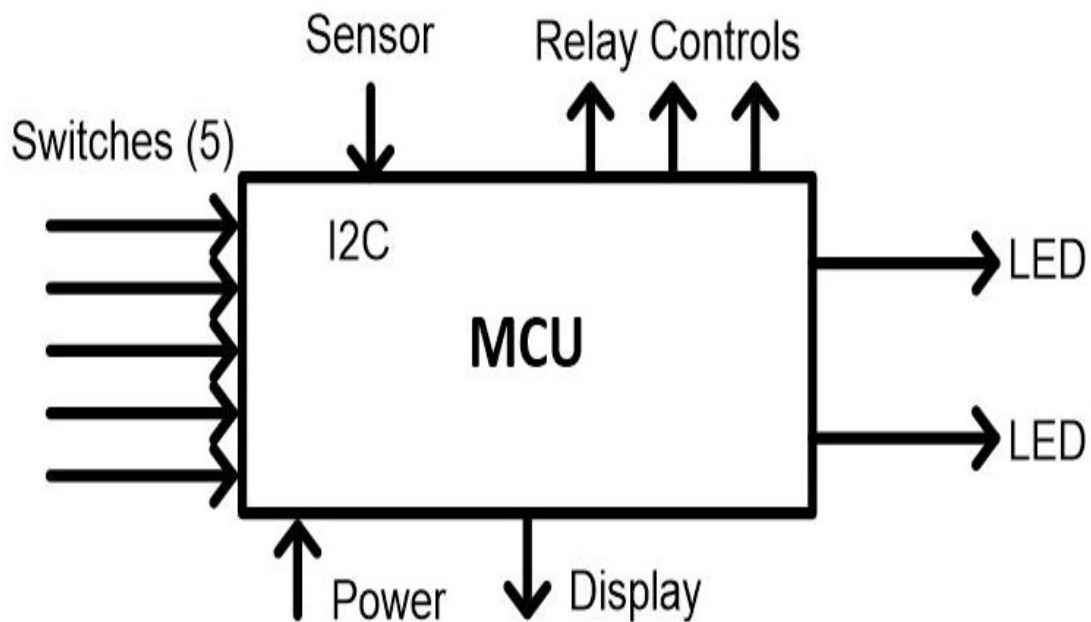


Figure 2-19. Single Controller Device

This example uses a generic MCU, without special internal electronics. Five input switches, an external sensor on a serial port, and several relay controls are attached. A serial port is set up to drive a liquid crystal display. This example is taken from an electronic thermostat, the sensor is for temperature, and the three relay controls are for heat, cool, and fan.

An industrial application illustrates multiple MCU's working together, but not electronically linked:

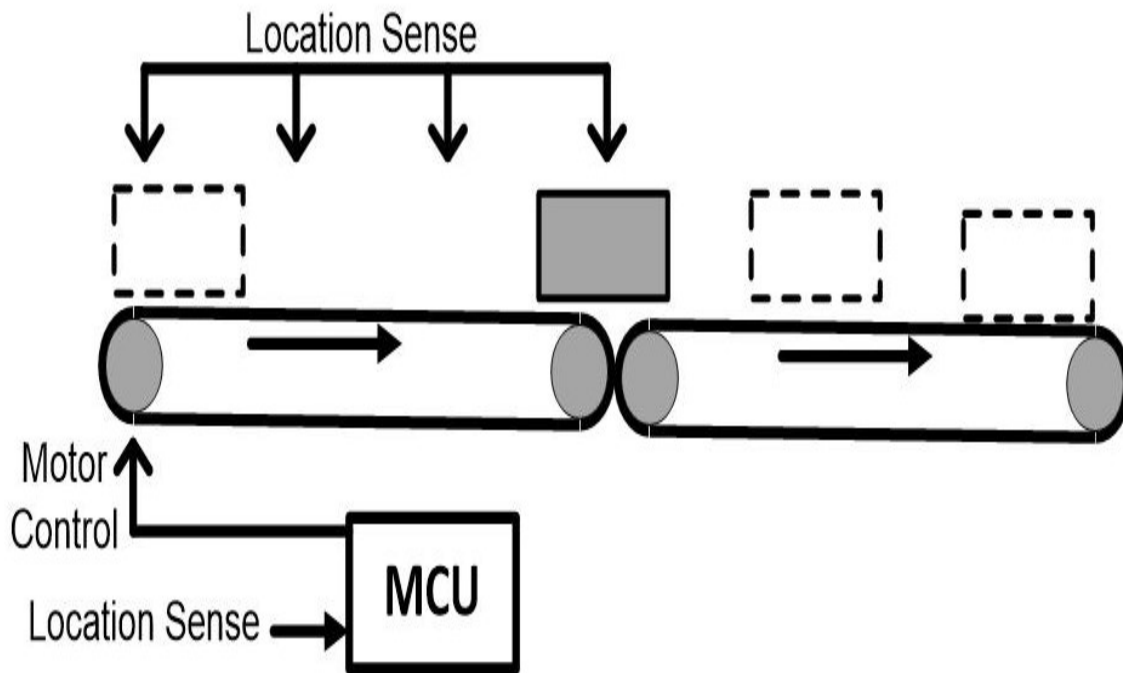


Figure 2-20. Multi Controller System Without Master

A conveyor belt system can detect the presence and location of packages along the belt. Each belt section has a separate controller and can turn the belt off when no packages are present (energy saving) and control the belt speed. If the package is not picked up by the next belt yet, the system can hold packages in the middle, waiting for the next belt to take up the package.

Communication and coordination between belts here is by the handoff of packages, rather than having adjacent belts in electronic communication. Setup and configuration here is minimal since each belt acts as a self sufficient system.

More complex industrial control systems or robotic devices frequently make use of a distributed controller scenario.

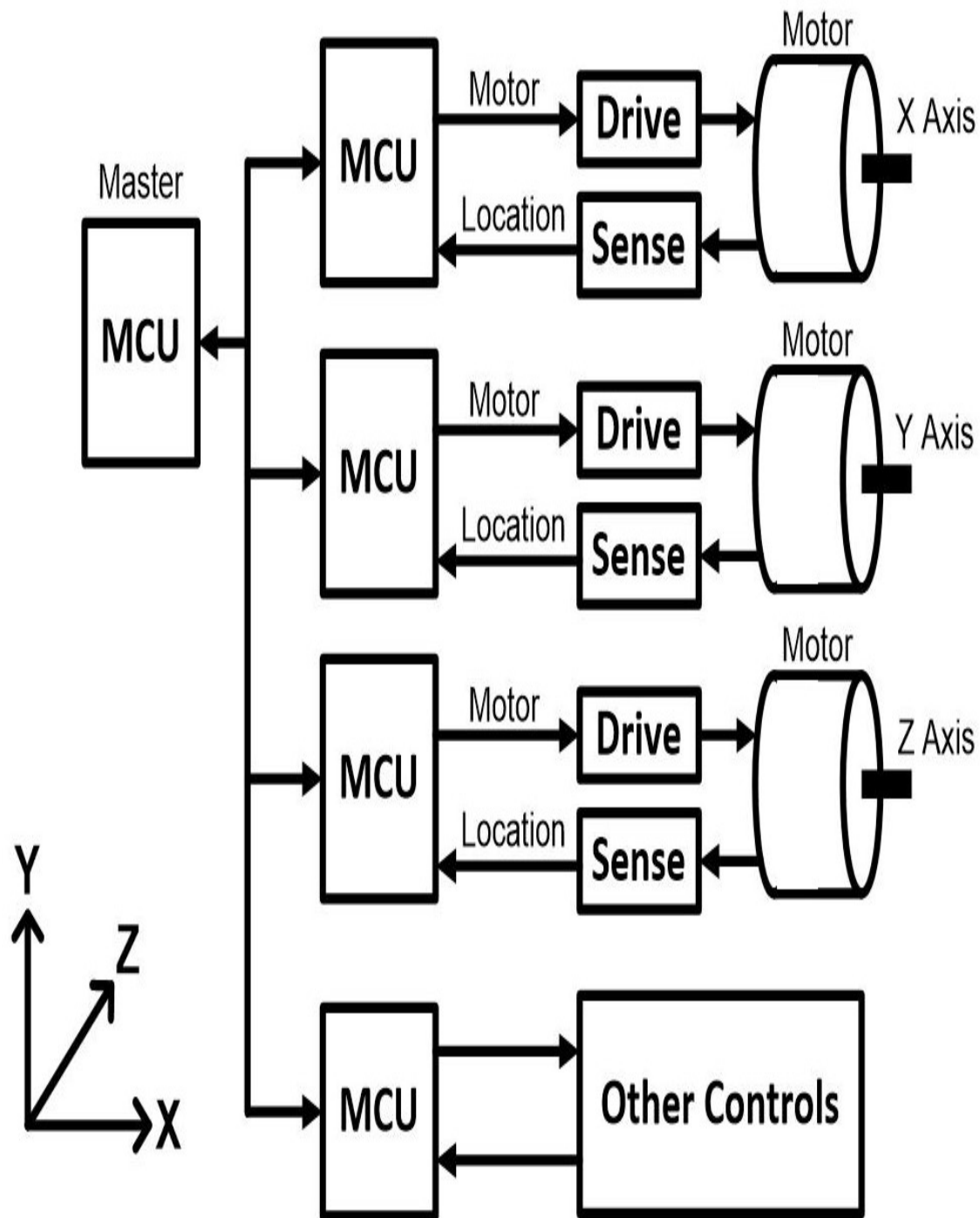


Figure 2-21. Multi Controller System With Sequence Master

In this example, a master controller runs the global system functionality, while local controllers deal with the detailed control and sense operations of

the device they are connected to. This is a common approach in industrial control, multi axis motion systems, robotics, and others. If communication is lost to the master MCU, local control MCU's can initiate a safety shutdown protocol. Frequently, this type of system uses the master MCU to download programming into the other MCU devices, allowing easy code updates to the entire system.

The examples so far include MCUs used to sense and control things. Frequently some form of data stream needs to be created from a device, and the resulting data needs to be processed.

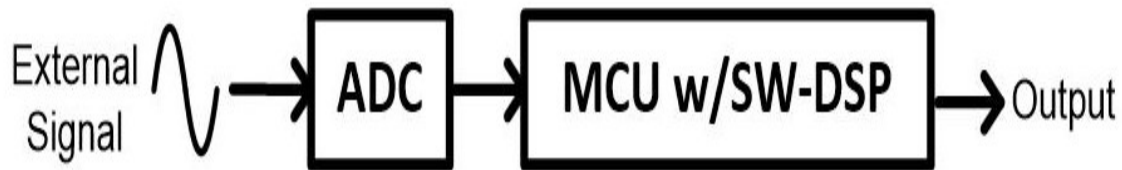


Figure 2-22. Streaming Data Processed Through Controller

If the data rates are suitable, an MCU can be used as part of the data streaming path. The viability of this method includes the data rate from the ADC, the complexity of the DSP algorithm, and the clocking rate of the MCU. Some configurations may choose to not stream the data through the MCU however:

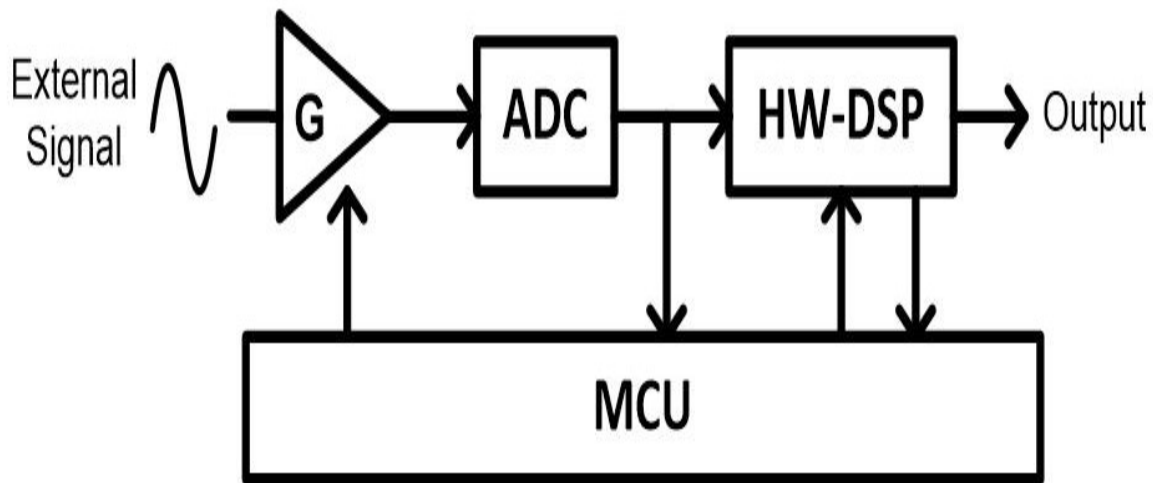


Figure 2-23. Streaming Data through Dedicated DSP

A common alternative is to stream data through a dedicated signal processing path, where the MCU controls the process, but is not in line with the data stream. Here, the MCU can sample the numbers seen at the ADC output, and adjust the gain amplifier to maintain an appropriate signal amplitude into the ADC. As well, the MCU can monitor the DSP system and make adjustments there as well.

This particular example is typical of a magnetic disk read channel, where the HW-DSP includes a phase equalization algorithm, bandwidth limiting LPF, and a Viterbi detection algorithm.

SOC devices frequently have MCU/DCU devices within them that are not even mentioned in the specification. This wireless microphone system is a good example.

Wireless Microphone and PA System

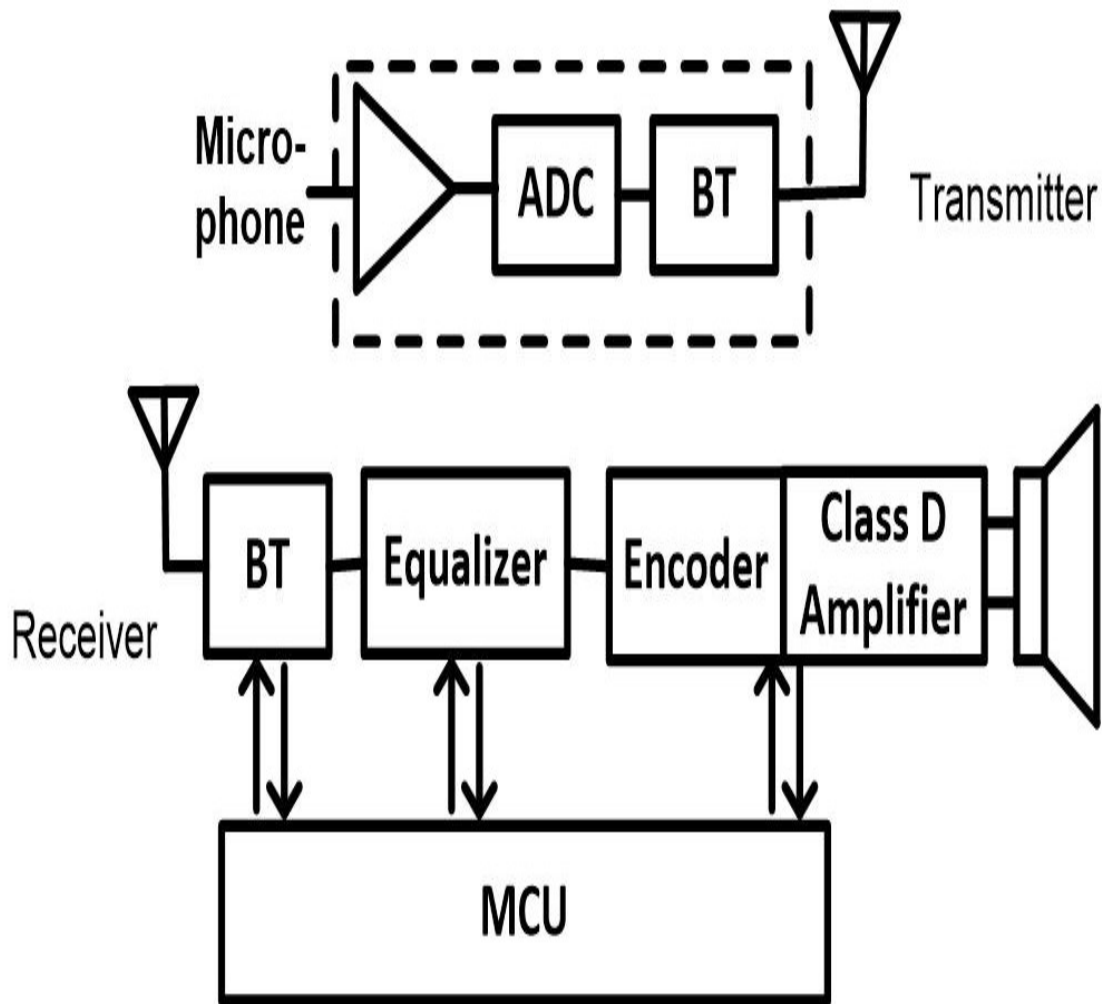


Figure 2-24. SOC Modules Used as a Set

The microphone has a DCU that controls the Bluetooth (BT) transmitter, and provides the control protocol associated with the BT standard. All the microphone electronics exist as a single SOC. The receiver has four DCU devices. One controls the Bluetooth receiver, one configures the equalizer, and one encodes the binary audio data to control patterns needed by the Class D audio amplifier. The prior three devices may actually be dedicated state machines within the SOC devices. The fourth device is an MCU that orchestrates the overall system.

A typical controller interface to peripherals is not just turning things on and off. The interconnection can have more complexity.

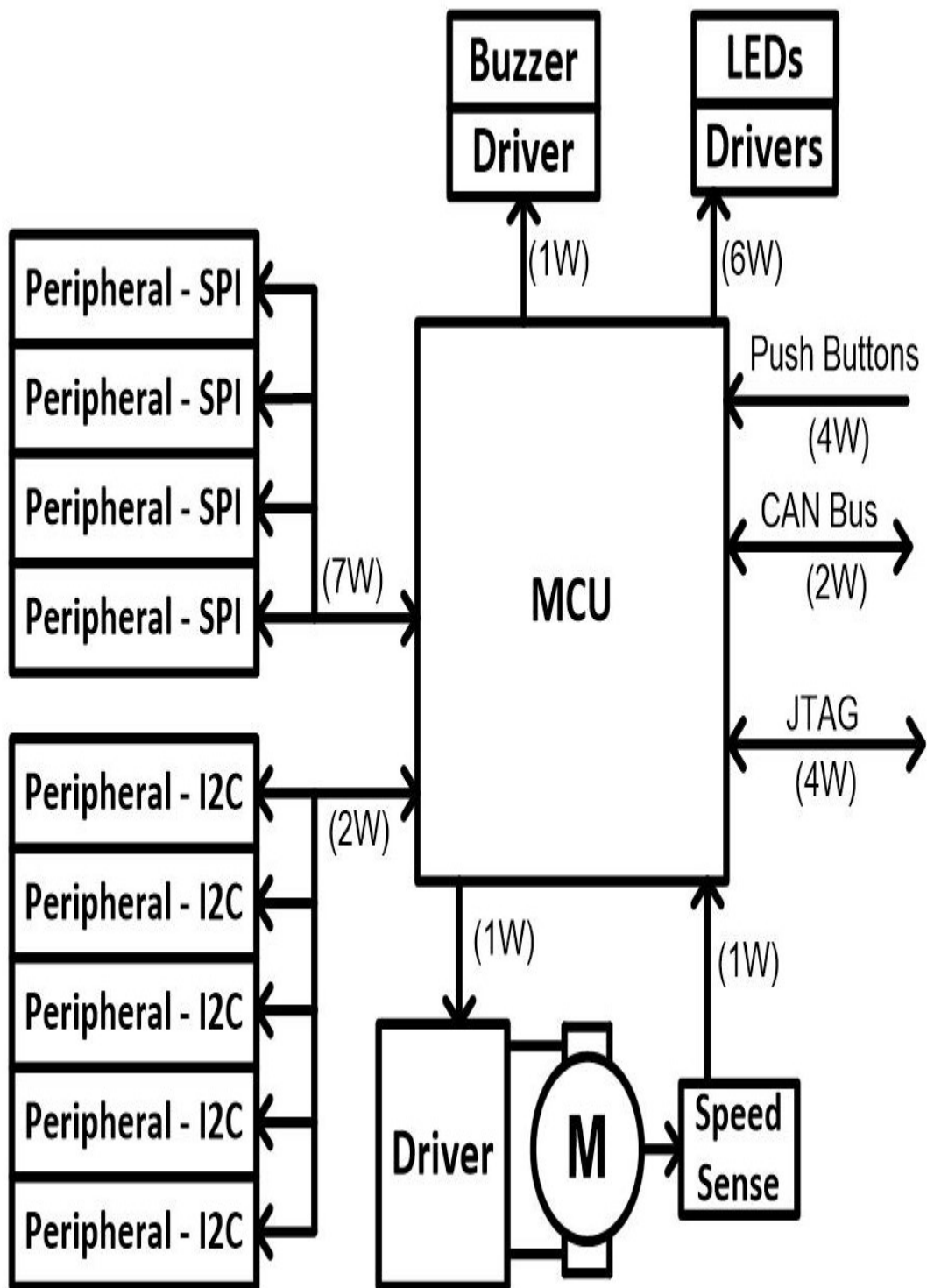


Figure 2-25. Typical MCU Interface To Peripherals

Adding some more detail to a typical controller, above shows the expected single input/output for switches, LED indicators, and buzzers. In addition, other external peripherals are connected through addressable serial ports, so multiple peripherals can use the same connections. Controlling a device often requires monitoring that device too. That idea is illustrated here with a drive circuit controlling a motor, and a sensor to monitor motor speed. Other useful external interfaces include a JTAG port for device programming, and a CAN bus to communicate off the board.

The examples presented here, give insight into deciding the global architecture of the system.

Determine Peripherals & Interconnects

The first step to picking a DCU strategy is to determine everything needing to be controlled or sensed. Some are obvious, but many items are often forgotten in an initial tally of peripheral interconnects.

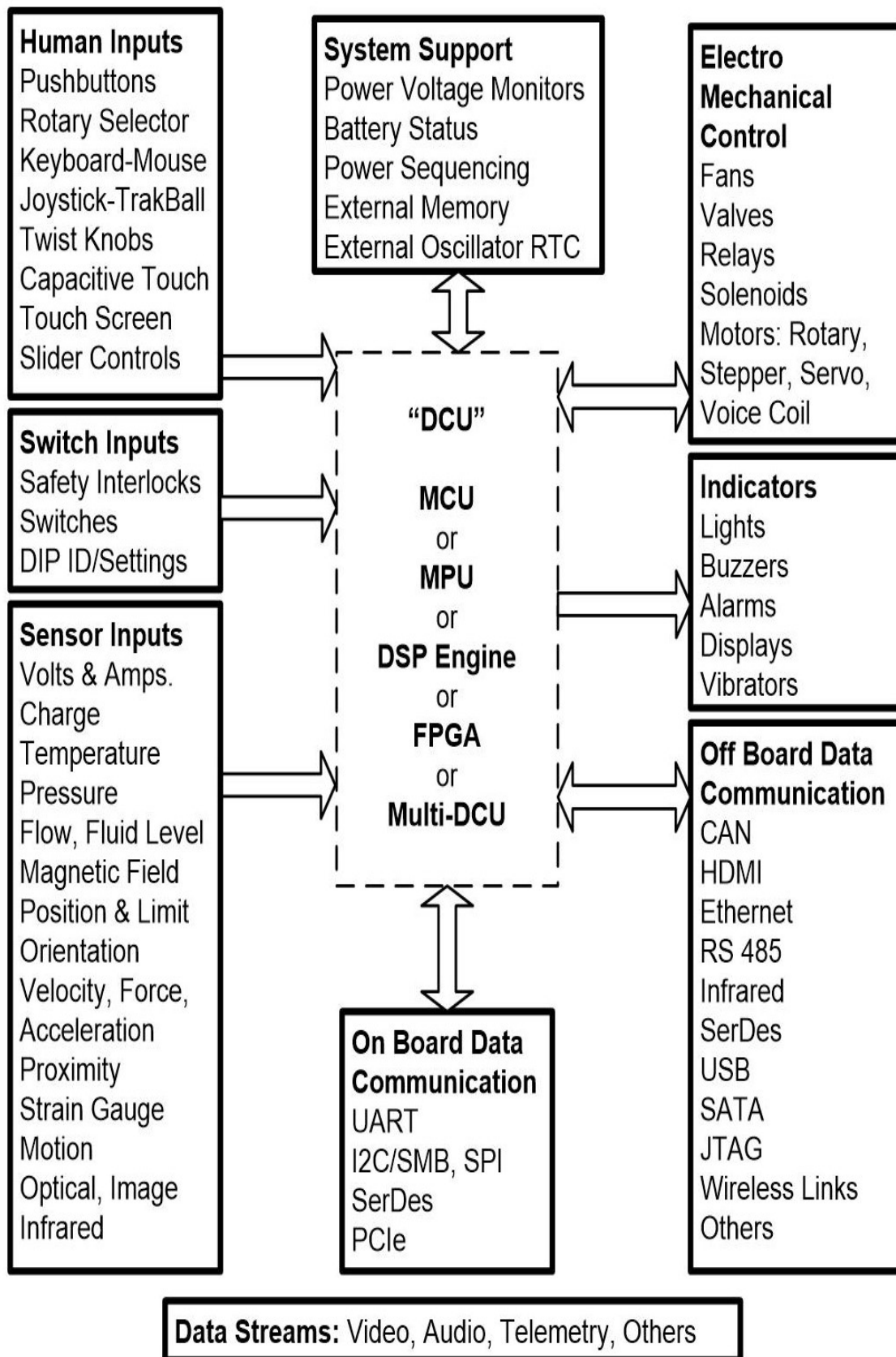


Figure 2-26. Control, Sense, Data In/Out

Looking at the possible peripherals by category:

Human Inputs

Most human inputs are simple pushbutton switches but there are exceptions. Most QWERTY keyboards interface through a USB port. Key pads are often done using a matrix switch structure, so that a 4X5 keypad uses 9 ports, instead of 20. Slider controls, twist knobs, and rotary selectors can be done using either an optical encoder or wafer switch, to give binary outputs, or a potentiometer, requiring an ADC to determine position. Determine specific human input devices and consult their specifications to determine an appropriate interconnect.

Switch Inputs

Switches are not limited to people pushing buttons. One example - Electro-mechanical systems use safety interlocks to stop active machines with safety covers off. Safety regulations frequently demand that a machine be rendered nonfunctional with exposed moving parts, and a safety strategy needs to be included. A safety interlock can include shutdowns for covers, latch solenoid to keep covers secured while things are moving, and motion sensors to determine when things have stopped moving.

DIP switches are still used to set a binary identity or configure a control parameter. However the availability of flash memory within the MCU allows most situations to be a reprogram of the MCU code.

Sensor Inputs:

Sensors come in many forms. A partial list is shown in the illustration. In addition to what the sensor is designed to detect, sensors can be divided by how their outputs provide information:

- **Threshold:** A threshold sensor is a device that changes state when the input exceeds a particular value. For the case of a pressure

sensor, it could be designed to change state at 35 PSI as part of dedicated module used for determining low pressure in tires.

- **Analog Output:** An analog output sensor changes voltage proportionally with the input. The voltage needs to be quantized to be useful to the MCU. A thermocouple is a common example. The typical thermocouple comes with a reference table linking temperature to output voltage. A suitable ADC will be needed to find the voltage, and interpolation between voltage and temperature sensed can be done with the MCU.
- **Pulsed Output:** A pulsed output sensor is often used for motion detection. Typically these devices emit a digital pulse for a fixed amount of linear or rotational motion. Using this information, position, velocity and acceleration can be determined by the MCU.
- **Data Output:** A data output sensor includes a fully integrated set of processing electronics and the sensor output is digital data from a serial port. Signal processing and an ADC have been embedded within the sensor body. This approach fits well with the mostly digital approach encouraged here.

On Board Data Communication

Communication on the PCB to peripheral devices is usually done using serial data ports. The commonly used methods at low data rates are I2C, SPI and older UART methods. Faster methods include PCIe and SerDes serial ports. (Details in digital communication chapter)

Off Board Data Communication

Within a multi MCU system, CAN bus and RS-485 are often employed. Between systems Ethernet is a reliable and suitable solution. Wireless methods (Bluetooth, WiFi, ZigBee, et. al.) are useful options. Depending on the specific application, other options are suitable. Further details are in the digital communication chapter.

Indicators

This includes “front panel” indicator lights, buzzers, alarms, vibrators, and displays that the user interacts with. Some displays will require a specialty port within the MCU.

In addition to front panel indicators, it is valuable to include indicators that aid debugging and troubleshooting. LEDs placed on the PCB can be used for in the field debug using binary error codes or a blinking “heartbeat monitor” showing that the system software is running. High failure rate items often benefit from a status indicator to tell a field technician of specific problems.

Electromechanical Controls

Fans, motors, solenoids, and servomechanisms all require interface driver circuitry to connect with the MCU. On-off control is generally sufficient for things like valves, relays, and solenoids. However, variable duty cycle drive capability, (aka Pulse Width Modulation, PWM) would be applicable for fans, rotary motors, and linear servos.

A speed sensor is commonly used to monitor fans and other rotary motors. A common strategy is to close a spindle speed control loop through the DCU. Linear servos and other voice coil motors require a location sensor to provide position information and location control SW generally implemented within the DCU.

Stepper motors require a driver circuit that provides both direction and step controls. Also, stepper motors need to be supported with a “home” location sensor.

System Support

The structure and control of the internal system requires support from the DCU as well. Some common features included are:

Battery State of Charge

A voltage monitor or a charge monitoring system is often developed for battery status or fuel gauge. This frequently needs a supporting ADC.

Power Sequencing

The DCU is configured with control lines to start-stop the system power supplies. This is important for electromechanical devices to allow orderly device control and avoid destructive behavior. The DCU power is always on, thus allowing live control when any power is present. The motor power is controlled by the DCU.

Power Monitoring

Voltages of the various internal power supplies can be connected to the DCU through an ADC to allow DCU monitoring of the supply voltages.

External Memory

May need a dedicated port interface if large amounts of data are being processed.

External Clock Source

A common strategy uses an external resonator “Real Time Clock” (RTC) reference at 32.768 KHz. This divides by 215 easily to create an accurate 1 Hz reference.

In order to make all of these peripherals function reliably with the DCU, most inputs/outputs will require some form of: SW signal processing, driver/receiver circuitry, ESD protection, or EMI limiting. Several later chapters of this book are dedicated to these issues and the details

thereof. With these examples, a list of all peripherals and their needed connections should start to emerge.

Data Streams

The processing of data streams needs careful attention because it is time restricted. Designers need to define data rates from all devices, including samples/sec, and bits/sample to quantify throughput needs. In addition, understanding the DSP methods to be used in processing each data stream will guide decisions on DSP HW & SW needs.

Avoiding Serial Communication (SerCom) Bottlenecks

SerCom ports (I2C, SPI et. al.) are frequently configured with multiple peripheral devices attached. As an example, a single I2C interface can have up to 128 devices attached.

In order to avoid traffic bottlenecks, an estimate of the worst case scenario for traffic through the port can be made. Determining the time period of a data transfer for each device, and how frequently each device is active. A summation of all device times gives an indicator of how “busy” the port is. No well defined rules exist here, but staying below 25% is suggested. Even at 25%, probability is such that for every 1 in 4 attempts to use the port, it will be busy servicing a different device. A higher percentage of use is possible with only the DCU requesting data rather than multiple uncoordinated devices trying to access the port.

Two other options can reduce SERCOM collisions. First, running the port at a higher data rate reduces the percentage of use. The I2C protocol has four different data rates in the specification. Second, MCUs are available that have multiple serial ports. Generally MCUs runs at a faster clock rate than the serial port, so the MCU can service multiple ports at the same time.

Direct Memory Access (DMA) for Data Transfer

A useful feature when moving large blocks of data in and out of the system is the capability to do Direct Memory Access (DMA). A DMA control feature has the capability to access internal memory and the internal data bus of a MCU, without CPU involvement.

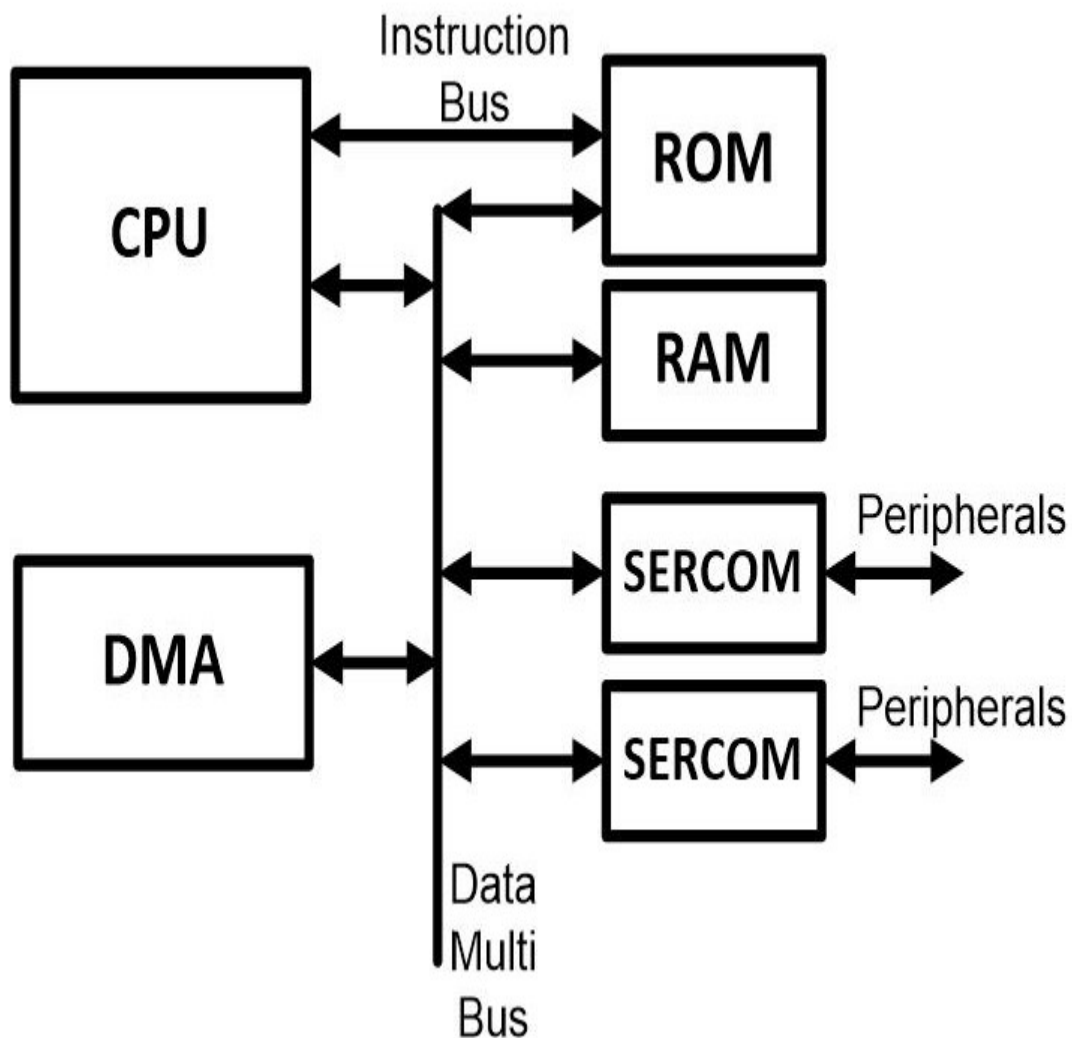


Figure 2-27. DMA Methods Used Between Memory Blocks and Peripherals

Here the DMA device can work independently from the internal CPU, and efficiently transfer data between memory blocks or memory and an external

port. The DMA device can transfer data with fewer clock cycles than using the CPU to read, then write, each data byte.

Determine DSP Methods

Until now, DSP has been treated as a black box. Selecting an appropriate DSP strategy requires a closer look. Either the DSP requires dedicated hardware, or, a software algorithm can suffice. Using a DSP implemented filter as an example:

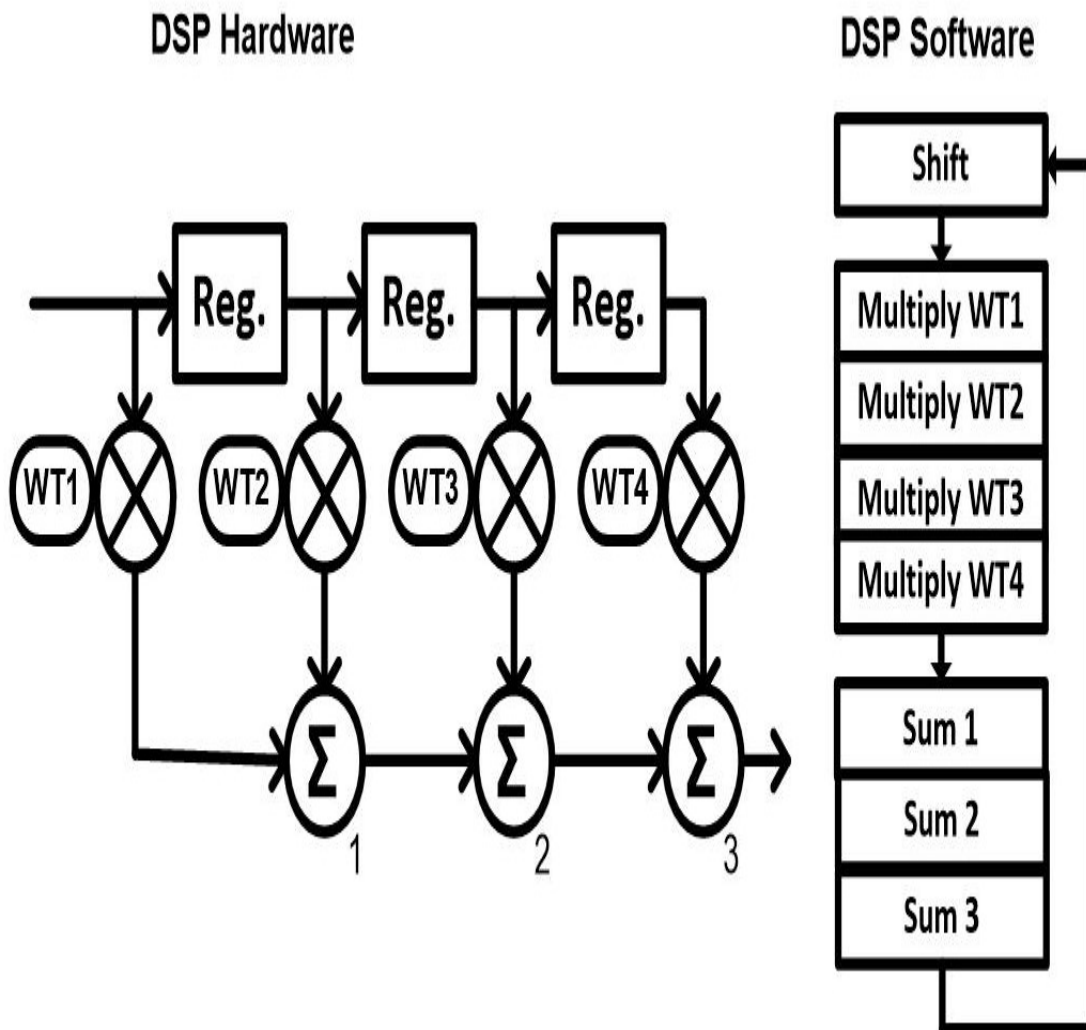


Figure 2-28. DSP FIR Filter, HW vs. SW Methods

This is the configuration of a 4 tap Finite Impulse Response (FIR) filter. The FIR filter on the left is a HW based implementation. Three shift registers, plus the input, hold 4 sequential samples of the signal. The 4 samples are multiplied by 4 different values (WT1- WT4) and then summed to create the output.

On the right, is flow chart of a SW implementation of the same FIR filter. The 4 sample values are held in a memory stack, each are multiplied by WT1- WT4 and summed, providing an output. Then the oldest sample is deleted and a new sample is put on the stack repeating the process.

Both HW and SW methods are commonly used. In determining which method to use, certain things come into play:

- Sampling rate of data stream
- DCU clock rate
- Complexity of DSP algorithm

Essentially, selecting SW vs. HW for the DSP is a question of how fast the data comes in vs. how quickly it can be processed. How “busy” the DCU is with other tasks needs to be considered also.

In addition, the delay time of the DSP may affect some systems:

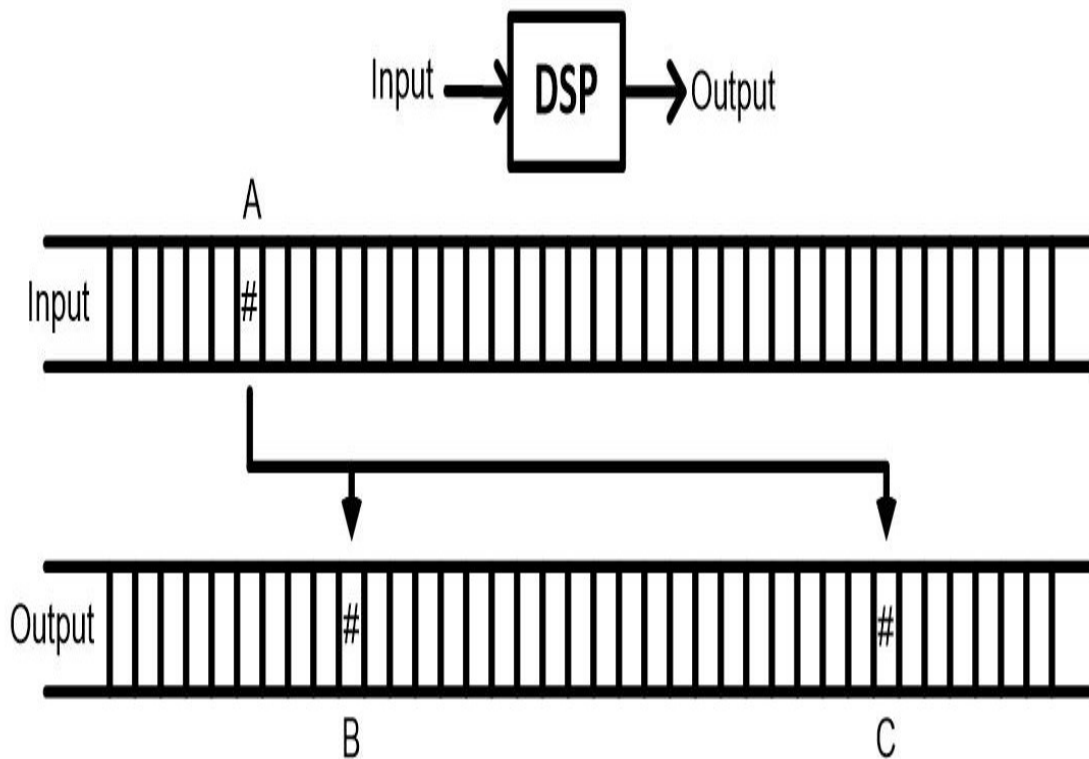


Figure 2-29. Data Stream Latency

DSP can be considered a pipeline. With an incoming data stream, and some math processing in the data pipeline, the output has a time lag from the input. That lag, is referred to as “time latency” or simply “latency.”

In certain situations latency can be an issue. However, in many cases it is invisible to the user. Two examples to consider:

Audio Latency

An audio path that has latency under 100-150mS is generally not a problem. Longer delays can be problematic in duplex communication (two people having a conversation over a phone).

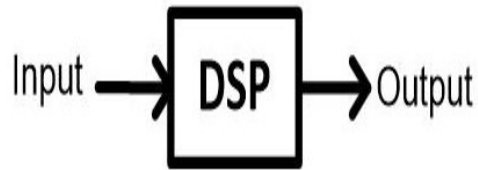
Control System Latency

Latency that exists within a control feedback system. If considered from the perspective of analog feedback, latency can be considered additive phase within a system. A digital system with too much additive phase in a feedback system can cause the system to become unstable.

Designers need to determine how much latency is tolerable, and accommodate that requirement with an appropriate DSP methodology.

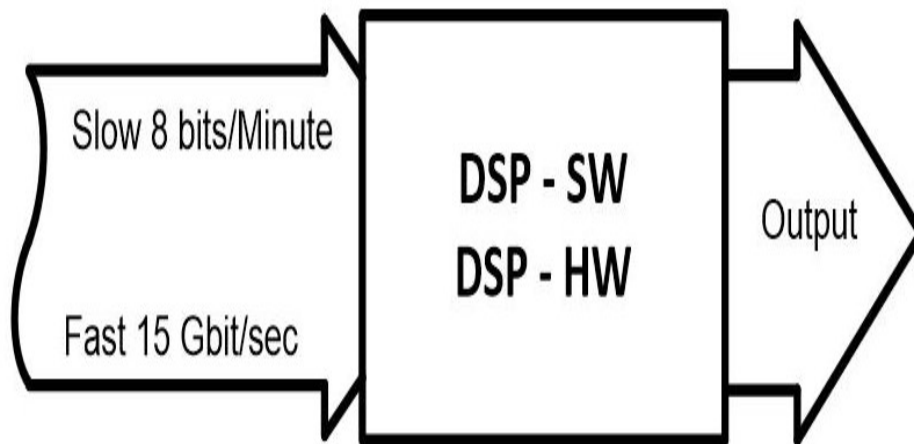
DSP Bottlenecks

A deeper look at two extreme examples of data rates can be useful. At low speeds, a thermal monitor needs to transfer an 8 bit number once a minute. At fast speeds, a high resolution video data stream, that uses 15GB/s coming from the camera sensors.



Slow Case: Thermal Monitor – 8 bits/Minute – SW DSP

Fast Case: RAW 4K HDTV – 15 Gbit/sec – 3X HW DSP



Demands for Processing Speed:

- Sampling Rate
- # Bits per Sample
- DSP Function Complexity
- Boolean or Floating Point
- Complex Numbers ($a + ib$)
- Resolution of Floating Point
- Depth of Array (# Taps)
- Restrictions on Latency

Options for Processing Speed:

- Clock Rate
- Internal Structure (8,16,32,64 bits)
- HW or SW methods
- Optimized HW
- CPU, GPU, FPGA or DSP Core
- Multi-Core
- RAM size and speed
- Tight Code
- Avoiding memory swaps

Figure 2-30. Data Processing Bottlenecks

Most data rate situations lie between these two extremes, but these two examples illustrate the issues. The slow data rate can easily be handled by a SW MCU system running at a low clock rate. The fast data rate is more challenging. The 15GB/s comes into the DSP already divided into Red-Green-Blue pixel data streams at 5GB/s each, which SerDes and PCIe interfaces can readily handle.

Once serial data has been input to the DCU, the pixel data is transformed to parallel data of 10 bits/pixel, for an internal data rate of 500MB/s, which is well within (2020 era) digital ASIC and FPGA capability. As well, many state of the art MPUs are capable of dealing with 30 bits of data at 500MHz. Be warned that some MCU devices are designed for lower (<100MHz) clock speeds, and could not process this data stream.

Creating demands for processing power and speed include:

Sampling Rate

How fast are the data updates coming?

Bits per Sample

A single bit is easier to process than a sample defined by 128 bits of data. Single data samples are easier to process than an array of samples, such as the data in a video image frame.

DSP Complexity

A DSP algorithm can be any math function, and the processing procedure can be simple or complex.

Complex Numbers

Data points that use a number value other than scalars, increases processing time.

Floating Point Resolution

Floating point math increases processor time for higher resolution results.

Size of Data Array

Multiple sequential samples are often used to determine a single output sample. This is illustrated with a FIR filter and how many “taps” exist in the filter structure. Image compression algorithms use multiple image frames to create each compressed frame.

Latency Restrictions

Is there a time restriction from input to output?

Improving DSP Speed

Things that improve DSP execution time include:

Clock Rates

Faster clocks yield faster results. But, remember that faster clock rates also create more power consumption. If there is no need for a maximum frequency clock, don't use it. Savings will be seen in both lower power consumption and less radiated EMI.

If faster clocking is needed, remember that most MPUs are faster than MCUs. Also, both MCU and MPU devices designed to operate using lower power supply voltage will generally be capable of faster clocking. If devices that use 5V or 3.3V power are not fast enough, investigate devices designed to function from 2.5V, 1.8V, or lower.

Internal Structure

MCU devices exist with 8, 16, 32, and 64 bit internal bus and CPU structures. MCUs with 8 or 16 bit structures are capable of Boolean decision making and control, but 32 bit devices are more suitable to systems that need floating point math. Many 8 & 16 bit devices were originally fabricated on older 5V CMOS processes, making them

slower. If using one of these devices, a modern functional equivalent is often available.

HW vs. SW methods

A dedicated HW solution (DSP engine, FPGA, ASIC) can be designed to run faster than a SW solution on a MCU platform.

Optimized HW

Many DSP engines have been marketed with structures optimized to the math functions common to most DSP functions.

Multi Core

If the function can be partitioned to run on multiple cores, a parallel processing approach can be used.

Tight Code

In the era of faster clocking rates and more powerful processors, improving SW efficiency is not that common. SW structure can change execution time significantly, especially for array processing and repetitive floating point operations.

Avoiding Memory Swaps

Most MCU systems will not need to deal with memory array swaps to mass storage. This issue is more common for a MPU running a multi-tasking OS. It takes time to move data and avoiding transfers of big blocks of data can expedite things.

Determine DCU Internal Features

With external connections determined and a better understanding of DSP needs, the next step is defining what internal features are needed.

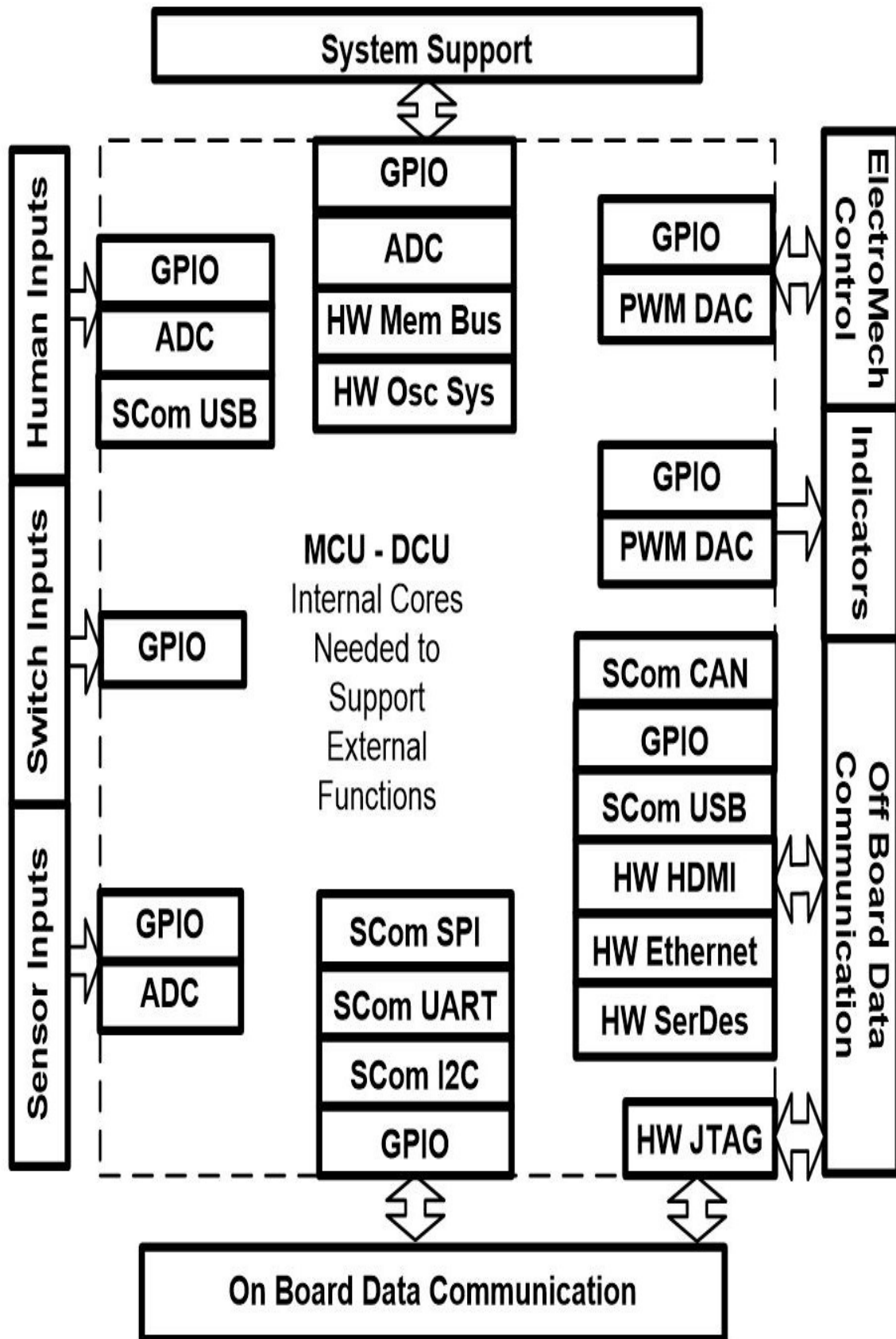


Figure 2-31. Determine needed controller cores and features

General Purpose Input Output Ports (GPIO)

From a list of peripheral devices, an estimate of the number of GPIO can be made. Capabilities of GPIO should also be determined, namely:

- Port enabled to service interrupt inputs
- Current sink-source capability
- Voltage high-low limits using internal drivers
- Vmax output capability using external pull-ups
- Vmax input tolerance when driven by external devices
- Programmable port configuration capability, including:
 - Output push-pull
 - Open drain with an external pullup
 - Selectable internal pullup
 - Drive strength (and speed) adjustable
 - Capability to tri-state the output
- For unconnected ports, how to render the pin inactive
- What internal functions can be routed through each GPIO (PWM output, Comparator Inputs ADC input, et. al.)

Serial Data Communication (SERCOM)

Serial communication ports include CAN, RS-232, RS-485, UART, USB, I2C, SPI, Ethernet, HDMI, memory card interfaces, PCIe, SATA, JTAG and others. Some SERCOM ports are generic, and can be configured for various purposes. Some require specialty internal HW to be compliant with the standard. Others may require external circuitry

components to make functional. All require some form of SW driver to properly respond. Check to see if the MCU vendor has a suitable SW driver.

Internal ADC Capability

Many MCUs offer built in ADC's. Check for a suitable number of bits and conversion rate. Some include programmable gain input amplifiers, which should be checked for noise, linearity and BW performance. Preferably, the ADC input is differential and not ground referenced. Differential inputs reject environmental noise better.

Internal DAC Capability

Many MCUs have PWM DACs available. Some will offer Sigma Delta DACs or other variants. Check for number of bits, conversion rate, and what external support components are needed.

Comparators

Voltage comparators are often useful in detecting a voltage transition over a trip point. Hysteresis capability or programmable voltage reference can be offered for some devices. Check specifications for input voltage limitations, response time, and input offset parameters.

Real Time Clock (RTC)

A RTC is for time based measurements. This is often implementing using a crystal at 32.768 KHz, which provides a 1 second clock after, dividing by 215 times. Also, calendar functions are available that will keep track of date-time-weekday data.

Clock Generation System

Most MCU devices have an internal oscillator to create a system clock. Many include a programmable clock multiplier, internally implemented with a PLL, to create a higher frequency clock. The basic clock frequency can be inaccurate when using an internal clock. Due to that,

many devices also support a crystal oscillator for a highly accurate clock frequency.

Power Consumption

Power consumed by the MCU will be determined by the MCU design and the clock frequency. Faster clocks consume more power.

Understanding the power consumption is especially important if the device is battery powered. There are many MCUs that have been optimized for low power consumption, and are defined as a “low power” product line. Specifics of power consumption need to be checked on a case by case basis.

Power Supply Voltage

As the CMOS transistor has shrunk, the maximum power supply voltage has also been reduced. Depending on what CMOS generation was used, the power supply may be 5V (older) to 0.9V (newer).

Different MCUs are fabricated on different generations of CMOS and a broad selection of devices exists for 2.5V, 3.3V and 5V power. Specialty devices with higher voltage interfaces also exist. Many “level shifter” chips are available to allow devices with different voltages to communicate with each other.

Programmable Memory

Programmable memory (aka Flash Memory) is implemented with some form of EEPROM that retains data when the power is off. There are different types of EEPROM and some are limited in the number of times they can be overwritten. If many (>1000) overwrites are expected, check this specification.

Needed memory size depends upon program complexity. Compiling the code for the device will give an estimate of needed memory size.

Lacking a complete code build, a safe start is to get the largest memory option for the MCU selected. As the project finalizes, memory size can be reduced as needs become better defined.

RAM Size

The majority of MCUs have RAM under 1MByte. Most embedded control systems don't need huge amounts of RAM. If large arrays of data are needed a MCP with an external RAM array may be more appropriate. Also, additional memory (both RAM and EEPROM), can be added to any MCU as a serial port connection.

Interrupt Handling Capability

Check to see which GPIO ports can handle interrupt inputs. Also review how the internal interrupt controller functions and processes an interrupt.

Floating Point Math Capability

Floating point math is best performed on a 32 bit (or more) processor. Although floating point routines exist for 8 bit devices, the HW limits performance. If a need for floating point math is expected check that the vendor has available a set of suitable code routines for the processor and that it all meets the IEEE-754 Standard for Floating Point Arithmetic.

Operating Temperature

Most MCUs will function between -10 C and 85 C. Devices can be found that will function down to -55 C and some up to 225C.

Power On Reset and Brown Out Detection Reset

How a MCU responds to power activation needs to be predictable and safe, especially when a MCU controls power equipment and motorized devices. Many newer MCUs include appropriate control circuits to safely assist with power cycling.

Radiation Hardened Designs

MCUs with structures that can withstand radiation ions are a specialty item for the military, satellites and aerospace applications. "Rad Hard"

devices are available from vendors that support the defense and aerospace sectors.

Direct Memory Access Capability

As described before, DMA can expedite the moving of data without involving the CPU.

Watch Dog Timer (WDT)

A WDT is a counter-timer used to detect if the system stops normal operation. If the system SW is operating properly the WDT counter is being cleared frequently. If the system freezes, the WDT counter is not being cleared, and the WDT is activated. Depending upon the specific features, a hardware reset occurs, or a special restart routine is called. Parameters of the WDT can be configured. A WDT is a useful safety feature if the code was not configured to self recover.

Timer-Counter Circuits (T/C)

Keeping track of time by counting clock cycles is the primary function of TCC. A TCC in a MCU can be used for:

- Event timing
- Creation of waveforms and PWM duty cycle controls
- Creation of delays
- Periodic execution of service routines
- Creating a slower clock

Multiple timers can be very useful. Check the features of TCC devices to see if they have sufficient bit range for your expected uses and capabilities to select both input and output connections.

Sleep and Low Power Modes

Some devices can shut down most of the chip functionality, until something requests a wake-up. Other devices can be configured to sleep for a defined period, then wake up, perform a check status and action routine, and then go back to sleep.

These methods are unique to each vendor. If the system functions in an active-inactive manner, this feature can be a valuable power savings technique especially for battery powered devices.

In sum, these features are available on various MCU devices, and the above should allow you to itemize a list of needed internal features.

Physical Package Considerations

Frequently the same DCU is offered in a multitude of package options. Newer options have gotten smaller while pushing designers to spend more time over a microscope to deal with the small form factor.

All 48 Pin Device Packages



Figure 2-32. Chip Packaging Size Comparison

Four of the most common 48 pin packages are show in a scalar size comparison. Some comments:

DIP (Dual Inline Package)

A through-hole device still available for some devices. The large package size is needed to support the wide pin spacing of 0.1" (2.54 mm) which creates a large amount of connection inductance in each pin. The DIP is not for modern commercial designs due to the large size and high frequency limitations.

QFP (Quad Flat Package)

A surface mount device, with pins easily accessed outside the perimeter of the body. This package is commonly used with external clocks under 300 MHz. Presently in widespread use, up to 256 pins.

QFN (Quad Flat No-Leads)

A surface mount device, without externally protruding leads. Contacts are at the perimeter of the package. By eliminating the protruding leads the inductance of the connections are reduced over QFP devices, and the package has better high frequency characteristics. QFN packages are used over 1 GHz with up to 352 pins.

Ball Grid Array (BGA)

A surface mount, with contacts arranged in a grid across the bottom of the package. Devices presently approach 3000 pins, and are very popular for high frequency and small form factor designs.

Many other package options also exist. These are common options, and illustrate the limits of size, pin count, and frequency.

Off Chip Features & Support

In addition to internal MCU features, items beyond the chip need to be considered. Ease of use, documentation quality, and anything that expedites development, are all important.

Integrated Design Environment (IDE)

Chip vendors provide SW tools to develop code, and HW to program the MCU. This usually includes a code editor, simulator, compiler-linker, interface for download, and a debug system. Vendors want their chips designed into products, so must provide a quality IDE, and make it free to download and experiment with.

Documentation

Some developer's manuals for complex MPU and MCU devices are thousands of pages long. Simpler devices are less so. Make sure that the material is sufficient in detail, well written, organized, and easily used.

Scalable Chip Family

Many MCUs are part of a family of chips with different number of GPIO, variations on internal features, and other options. These devices try to maintain both software and design tool compatibility across the chip family. This allows designers to change between chips without a total redesign, or learning new design tools.

Software Libraries and Application Programs

Does the vendor provide properly documented and debugged support software? As an example, downloading a SW control protocol for any serial port structure is a valuable time saver. If a vendor provides a particular port interface or internal feature, check that support code for it is also available.

Support Community

Many user groups with online support and discussion forums exist for MCU product lines. Some are independent of the vendor, and others are

vendor supported. Having access to designers that use a family of chips and IDE tools is invaluable when you need a question answered quickly.

Evaluation Boards and Development Kits

Vendors frequently produce evaluation boards and other development kits to support products. An evaluation board allows a quick and low cost way to start running code on an actual device. A common strategy uses an evaluation board for code development while a PCB is being created, giving a few weeks of head start time.

Cost and Source Availability

MCUs are a commodity and many viable solutions exist. However, plug in compatibility between vendors is not generally available.

Consequently, a reliable vendor that can provide the needed quantities when your product ships, is important. Also, device cost may be counterintuitive. Some modern 32 bit MCU's cost less than 8/16 bit devices using older CMOS technology. Stiff competition in a flooded vendor field has resulted in lower cost solutions with more features. Check the cost for volume purchases, and don't assume fewer bits, or features, means lower cost.

Pulling it all Together

At this point enough information has been presented here to:

- Understand the “mostly digital” design approach
- Define your system architecture (single MCU, multi MCU w/wo master)
- Determine your peripheral devices (buttons, switches, sensors)
- Itemize your peripheral control connections (GPIO, DAC, etc.)
- Itemize your devices to sense (GPIO, ADC, etc.)

- Define what devices need both control and sense connections (motors, servos)
- Itemize devices connected to serial ports. (I2C, SPI, UART)
- Define off board communication needs (CAN, RS-485, Ethernet, etc.)
- Understand the time demands of data streams and their DSP
- Define where DSP needs enhanced processing or dedicated HW
- Decide between an MCU (SW) or FPGA/CPLD/ASIC (HW) approach
- Avoid SERCOM bottlenecks
- Determine what internal features are needed for MCU

Many of these items require appropriate drive-sense circuits to interface with devices. Detailed coverage of those methods are in following chapters.

Pick a DCU Configuration, Picking your MCU/MPU

Suggesting how to pick a device can be broken down to a several considerations:

Specialized Niche Function or Feature

If a MCU with some very specialized form factor or feature is needed the vendors may be limited. Some specialty functions on the same chip with a MCU may not be available from many sources. Frequently, that specialized function can be an external peripheral to the MCU, but getting it all on the same chip saves space, money and power.

Multi MCU Systems

The author strongly suggests that these systems use the same MCU in all locations. Although each device may have a different function there

will be a large amount of overlap in design tools and support coding. This will save both design and development time.

An alternative approach uses the same MCU everywhere except the master. If the master needs PC type features (video interface, keyboard/mouse input, run an OS) a MPU is often a better fit. A single board computer, with a CAN bus or RS-485 bus out to multiple MCU satellite boards is a common solution here.

General Use MCU Systems

For most systems, many vendors will have multiple devices that will get the job done. The first priority should be the “Off Chip Features & Support” considerations, namely:

- Does the vendor have a product line of suitable devices?
- Is the IDE easy to use and free of bugs?
- What’s available for ready to go SW and drivers?
- Is product documentation well organized and solutions are easy to find?
- Does the online support community show promise?
- What’s available for evaluation boards and development kits?

The strategy suggested is to pick a manufacturer and their product line first, then select a specific MCU. Frequently, ease of use and quality support are more important than a specific device selection.

Picking a Specific MCU

After narrowing the scope to a vendor, the lists of ports and necessary features itemized earlier defines a suitable device. Allow enough capability to support design changes. Include extra ports, extra memory, and processing capability.

Design details change, and desired product features invariably grow, so using a MCU with some excess capability is a good strategy for a first generation build. With a fully functional design, a cost optimized design can be pursued in a second generation build that has fewer ports and less memory. With a MCU that is part of a product line, (all using the same internal CPU) this allows changes with minimal difficulty.

The “Further Reading” section at the end of this chapter has a list of some of the more popular vendors of MCUs and component distributor search tools for specific devices.

Essential Concepts – Conclusions and summary

This chapter is about “big picture” ideas and global approaches to designing an electronic system. Some of the important ideas in this chapter are:

- Modern design strategies try to minimize analog circuitry
- DSP methods are versatile and widely applicable
- HW based control uses CPLD, FPGA and ASIC devices
- SW based control uses MCU and MPU devices
- DSP can be HW or SW based
- Chip sets or single SOC solutions exist for most high volume products
- Many system architecture options exist, with single and multiple DCUs

With that knowledge, a designer should be able to:

- Determine system architecture at the box level
- Itemize a peripheral list and needed interconnects

- Evaluate SerCom requirements and avoid bottlenecks
- Determine needed DCU features
- Evaluate an IDE and support environment
- Select a suitable DCU

In addition, designers should now have a suitable knowledge base to navigate the confusing maze of device specifications and feature descriptions of various processors. Remember that there are probably multiple devices suitable to the design project at hand. Having more than one correct answer can make the decision easier, although confusing nonetheless.

Further Reading

- Audio Latency: Section 7.4 of International Telecommunications Union G.107 (06/2015) ITU-T Series G
- Floating Point Math in a Microprocessor: IEEE-754 Standard for Floating Point Arithmetic
- Simulate or Build: “Factor Time and Money before Using a Simulator,” Jerry Twomey, May 03, 2014, Electronic Design
- MCU/MPU Manufacturers:
 - Analog Devices
 - Atmel
 - Cypress Semiconductor
 - Freescale Semiconductor
 - Infineon Technologies
 - Intel

- Maxim
- Microchip Technology
- Motorola
- Nuvoton Technology
- NXP Semiconductors
- Renesas Electronics
- Rohm (Lapis) Semiconductor
- Silicon Labs
- STMicroelectronics
- Texas Instruments
- Toshiba Semiconductor
- XMOS
- Zilog
- Component Search Tools:
 - Digi-Key Electronics <https://www.digikey.com>
 - Mouser Electronics <https://www.mouser.com/>
 - Allied Electronics <https://www.alliedelec.com/>
 - Jameco Electronics <https://www.jameco.com/>
 - Arrow ECS <https://www.arrow.com/>
 - Newark <https://www.newark.co>
- Useful web sites and design forums:
 - Texas Instruments Forums (VS): e2e.ti.com

- Microchip Discussion Forums (VS): microchip.com/forums
- All About Circuits (VI): allaboutcircuits.com
- Electro Tech Online (VI): electro-tech-online.com
- Arduino (VS): arduino.cc
- Raspberry Pi (VS): raspberrypi.org
- 6502 Specific User Group (VI): 6502.org
- Microcontroller (VI): microcontroller.com
- Atmel, AVR and Atmel-ARM (VS): community.atmel.com
- EDABoard, electronics forum for (VI): edaboard.com
- Digi Key Engineer Resources: digikkey.com/en/engineering-resource-center

NOTE

VI = Vendor Independent

VS = Vendor Supported

All of these sites are useful to designers. The VI sites will tend to be wide ranging on topics and the VS sites tend to focus on the vendors products. Search engines will point out many other sources of course.

Chapter 3. Robust Digital Communication

A NOTE FOR EARLY RELEASE READERS

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

This will be the 3rd chapter of the final book.

If you have comments about how we might improve the content and/or examples in this book, or if you notice missing material within this chapter, please reach out to the editor at arufino@oreilly.com.

Digital communication is the backbone of any system. Low frequency, short transit, data across a PCB is usually straightforward. But, when data moves off a board, goes a distance or encounters electromagnetic interference, challenges arise. Also, reliable data at high speeds requires special attention to signaling methods, transmission lines, error handling, and clock/data phasing. Countless systems work on the bench, and then fail in the field due to environmental challenges. Properly designed systems must survive real world EMI, ESD, variable power, and noisy grounds.

This isn’t a chapter on data communications theory. The intention is to give an understanding of widely used methods, so readers can make informed decisions of what interface to use or where they need a transmission line. Limited coverage is also given to older methods that designers should be aware of.

Commentary here is about CMOS logic. Special considerations for other logic families (RTL, TTL, ECL, etc.) are not covered. Modern methods put

logic functions in controller software or a single FPGA/CPLD. (Nobody uses a PCB full of discrete NAND gates anymore.) Multiple high speed data paths are now inside the chip, not across the PCB. Consequently most PCB digital communication now falls into two categories: low data rate setup and configuration control of chips, or high speed data pipelines between devices. Figures here may use single gates for illustration, but items here are applicable to the I/O ports of larger FPGA/CPLD devices which are invariably CMOS. Signals off the board or out of the system are special situations that also need to be dealt with.

We're going to cover a broad swath of digital communication topics here including:

- Ground referenced vs. differential signals
- Point to point vs. multi point networks
- Lumped vs. distributed networks
- When to use Transmission Lines with Termination (TLT)
- Clock distribution methods and timing skews
- Parallel vs. serial data interface
- Common serial ports and network methods
- Wireless data link methods

Designers can use this information to make proper choices for system communication methods.

Digital Signals, Physical Considerations and Connections

Before looking at suitable data interfaces, some background on how connections are made, what types of signals to use, when to use terminated transmission lines, and clock distribution methods need to be examined.

Limitations of Ground Referenced Digital Signals

Digital data is subject to signal integrity issues. Noise, amplitude loss, transmission line reflections, impedance mismatch, and others all contribute to the problem. Issues tend to grow as connections get longer, or data rates go up.

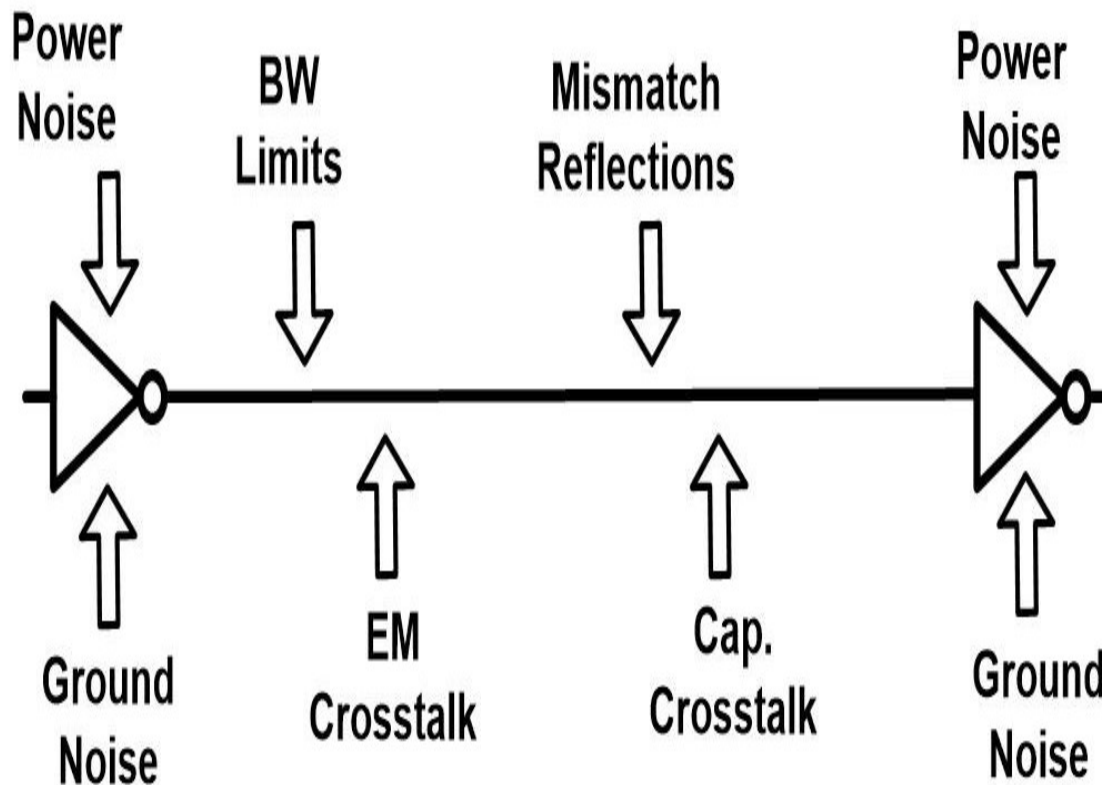


Figure 3-1. Digital Signal Corruption

A digital signal can be corrupted by noise on power and ground (P&G) at both transmit and receive sides of a data connection (Figure 3-1). Also, bandwidth (BW) limitations of the connection will cause higher frequency parts of the signal to be degenerated in amplitude and distorted in phase. Close proximity adjacent signals contribute crosstalk noise by capacitive and electromagnetic methods. Therefore, the rise/fall edges of digital signals make the impedance mismatches of a connection very visible. A connection with fast signal transients, and lacking a properly terminated transmission line, creates signal reflections further corrupting the signal.

Low Voltage Differential Signaling (LVDS)

Digital drivers can be ground referenced or use differential signals. Frequently differential signals use lower voltages than rail-to-rail ground referenced logic, thus the LVDS nomenclature. A standard (TIA/EIA-644) has adopted the LVDS name, however this text refers to all differential signaling using non rail to rail signals as LVDS.

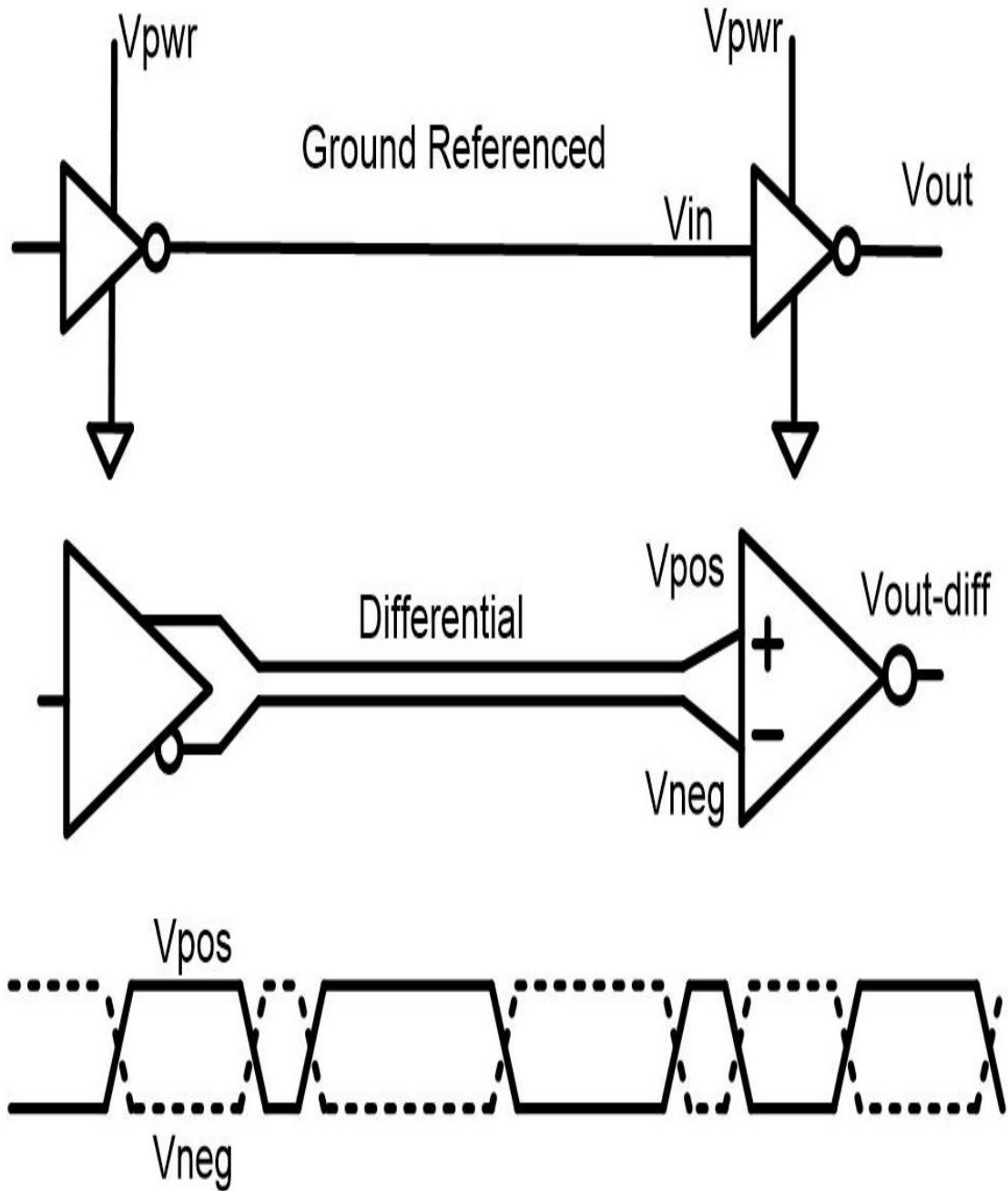


Figure 3-2. Differential Signals vs Ground Referenced

A ground referenced digital signal is dependent upon both P&G to create and detect a signal (V_{in}). P&G noise shows up on both the transmitted signal and also dynamically shifts the switching point of the receiver. A differential signal is defined as the difference, ($V_{pos}-V_{neg}$), between signals and thus minimizes sensitivity to P&G noise.

LVDS signals can be connected through either a differential strip line (PCB) or a twisted pair wire set (cable). Signals are equal and opposite in polarity.

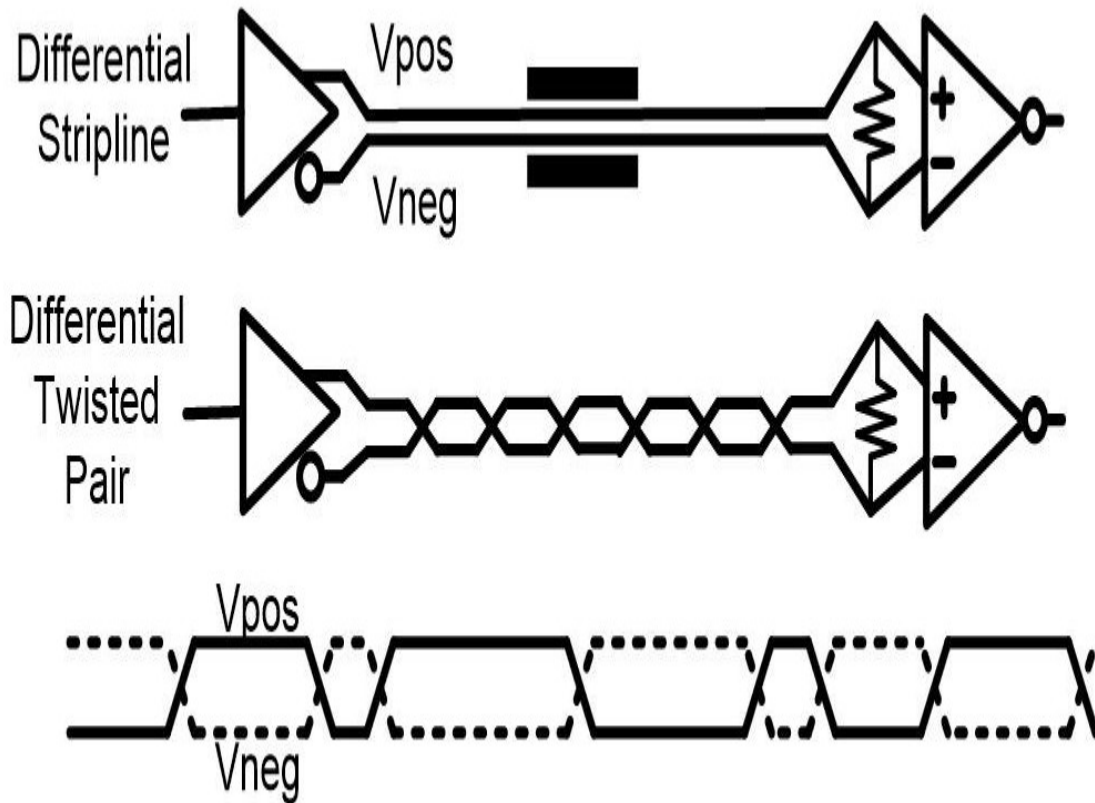


Figure 3-3. LVDS through Matched Differential Paths

Since the reception of LVDS is done as the difference between the two signals there are numerous advantages:

Less Noise Sensitivity

Differential signals are closely routed together making noise coupling common to both. A differential receiver rejects common mode signals.

Less Sensitivity to Power Noise

P&G noise is common mode to both signals, and also rejected by the receiver.

Lower Timing Jitter

Accuracy of the switching point is more consistent due to both common mode noise not contributing to time spreading of the signal, and a faster dV/dT of the differential signal. Due to being differential, the effective dV/dT enjoys a 2X improvement.

Less Bandwidth Sensitivity

BW limitations can cause attenuation of a ground referenced signal and create asymmetry in the received output. BW limitations of a differential signal will attenuate the signal, but LVDS receivers can detect a smaller signal without creating asymmetry.

Lower Power

LVDS commonly functions with 300 mVpp for each signal. At high frequencies, two signals, with lower voltage swing, frequently require less power than one rail to rail ground referenced signal.

Less Radiated EMI

With equal and opposite currents in two closely placed signals, EMI is greatly reduced. From a distance, the electromagnetic fields cancel each other out. Switched current surges in ground referenced logic acts as a transmitter, with the connection creating a transmit antenna.

For these reasons, differential signaling and LVDS methods have been widely adopted for both high speed and long distance digital connections. Ethernet, USB, PCI Express, CAN bus, among others, use LVDS methods.

Organizing Interconnects for Speed and Signal Integrity

With slow (<10 MB/s) data rates, soft (>3 ns) rise/fall times, and short (<25 cm) connections, getting digital data across a PCB is generally not a problem. The Slow, Soft and Short Net (SSSN) is easy to deal with, and makes up a lot of the digital connections in use. A more careful approach for higher data rates will be outlined shortly. Before that, understanding what makes a good network connection needs to be examined.

Any electrical connection, without controlled impedance, and with multiple connections at multiple locations is not friendly to high bandwidth signals. Disorganized connections can be thought of as a mixed network of RLC elements and unterminated transmission lines without consistent impedance. High frequency signals will quickly exhibit problems in such connections.

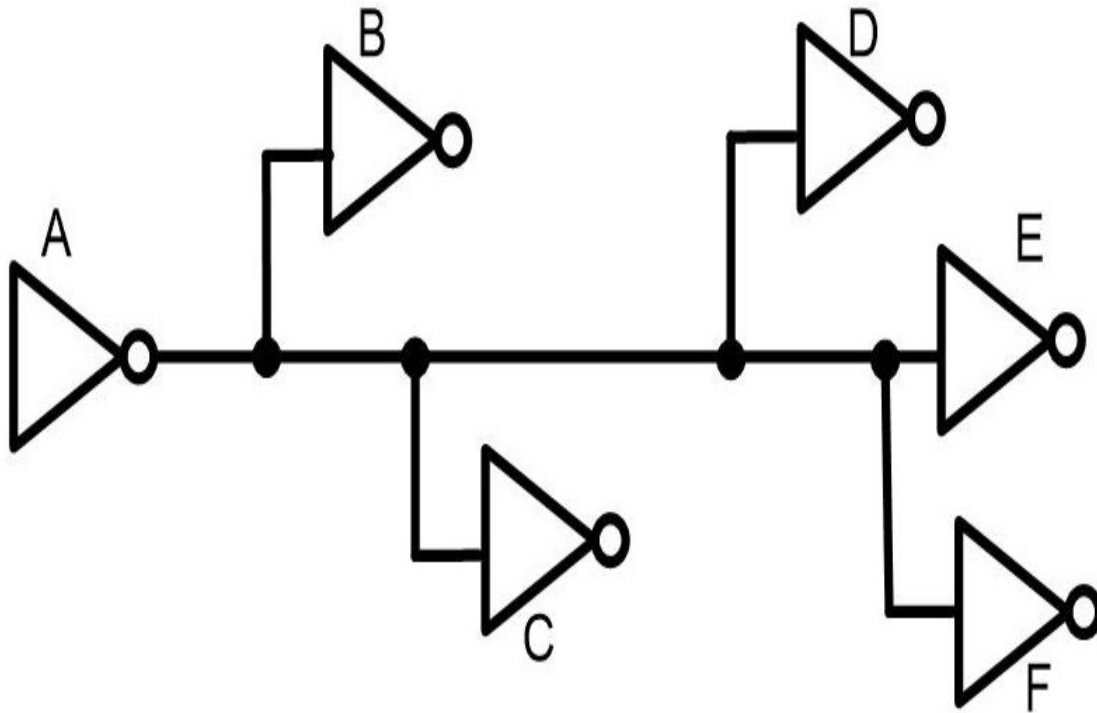


Figure 3-4. Problems with Multi Point Networks

This example shows a driver (A) that is connected to multiple load (B,C,D,E,F) with no well defined placement or routing. Consequently, the possibility of transmission line reflections, impedance mismatches and signal integrity problems abound here, especially when the data rates get pushed up. It is difficult to define what is wrong, because connection lengths and data rates are unknown. If it is a SSSN, it might work if device A can drive 5 loads and their associated RLC interconnects. Better definitions are needed.

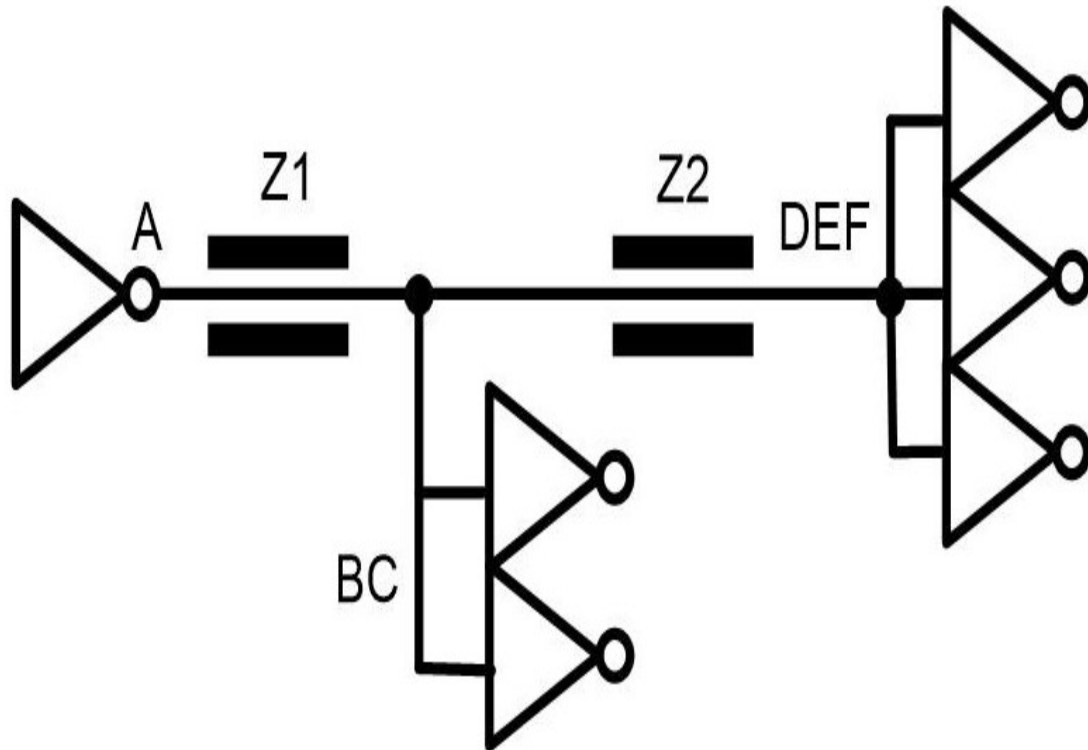


Figure 3-5. Reducing Transmission Reflections

Devices here are in two tightly placed target groups so you can treat groups (BC) and (DEF) as two lumped nodes. (More on this shortly.) This minimizes the undefined transmission connections. However, this structure still lacks a termination resistance, and includes a signal stub in the middle of a transmission line which can cause impedance mismatches and local line reflections as high frequency signals pass through. The (BC) group is referred to as a “stub connection” as seen in multi drop bus connections. Stub connections should be avoided or kept as short as possible.

Ideally, singular point to point transmission is the preferred method.

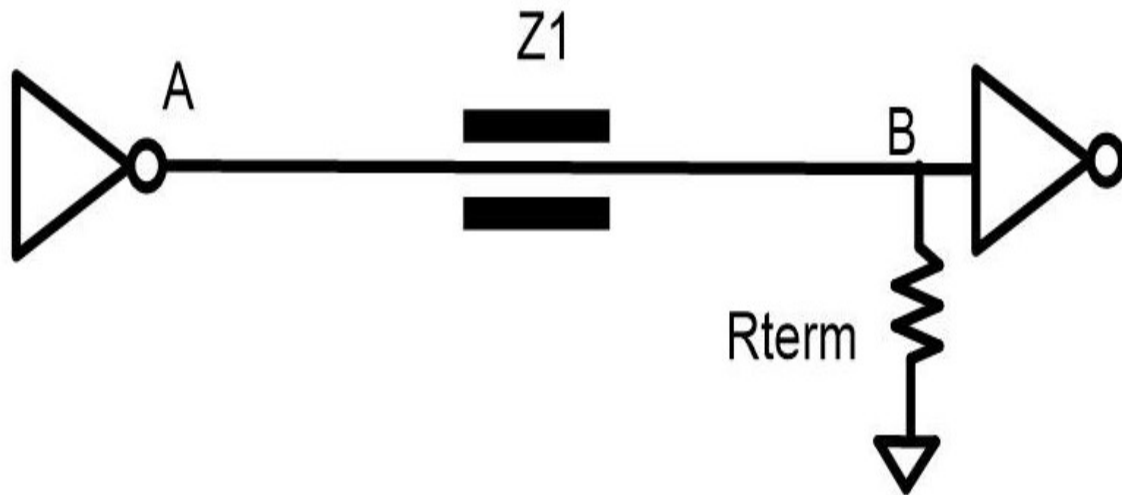


Figure 3-6. Point to Point Transmission

The above uses a controlled impedance layout for the connection and a termination resistance at a single receiver. For high frequency signals this minimizes line reflections. A singular point to point connection does not serve for all situations, but multi drivers can provide a solution.

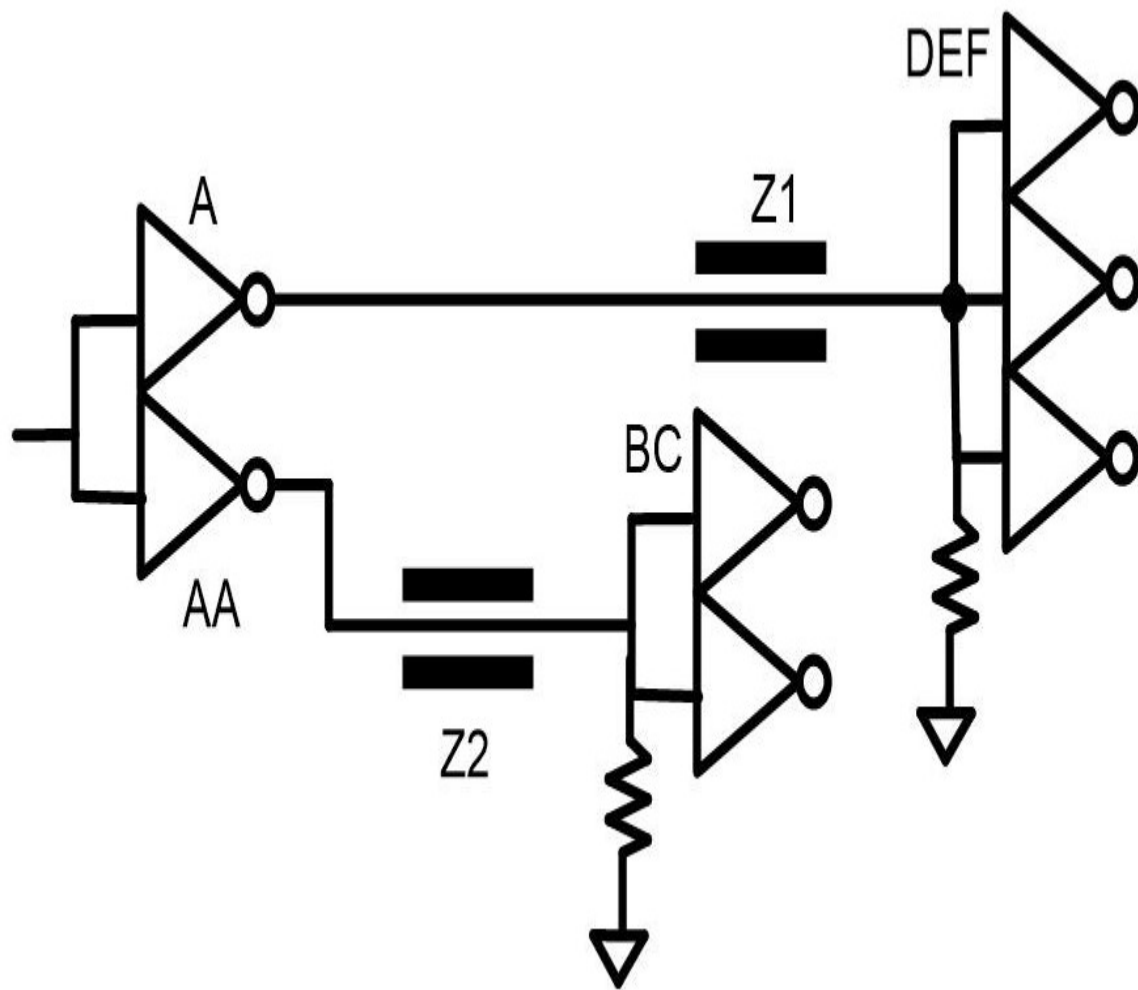


Figure 3-7. Point to Point with Multiple Targets

The above configuration provides a single transmission line between (A) and (DEF) as well as a single line between (AA) and (BC). Using separate drivers for each transmission line minimizes interaction between the two. Loads (DEF) and (BC) are each closely placed together and can be treated as a lumped network. Next, we'll discuss when to treat a digital connection as a lumped or distributed network and where transmission lines and terminations are needed.

Lumped vs. Distributed Networks

It takes time for a signal to travel down a wire, or across a PCB. A lumped network is where all connected components see the applied signal at effectively the same time. A distributed network is one where transmission delays cause applied signals to occur at significantly different times to the attached devices. Determining which way to treat a connection becomes a function of two items: the connection distance, and the transient signal applied. For analog and linear systems, phase as a function of frequency is used. For digital systems, the rise/fall time of the signal vs. connection length is useful.

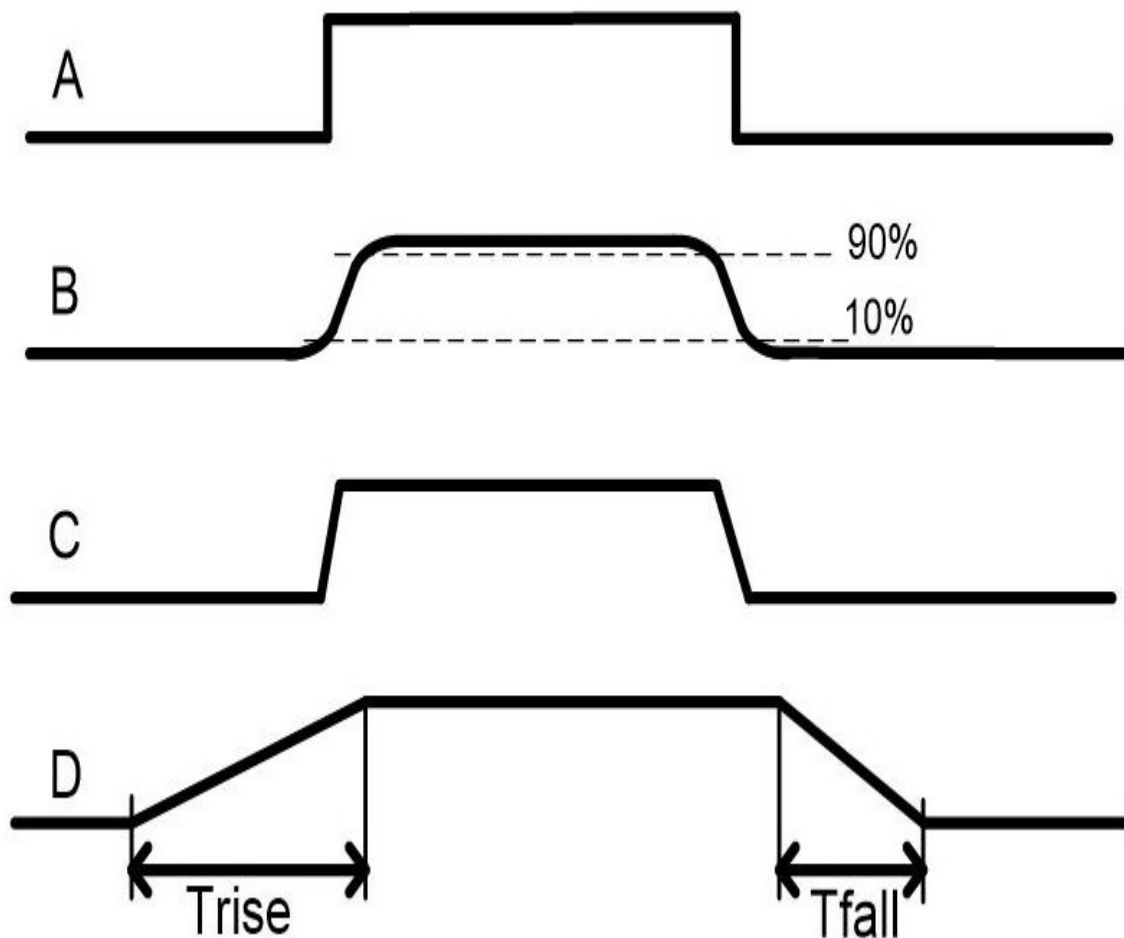


Figure 3-8. Digital Signal Rise Time

Ideally, signal (A) is shown with zero rise time. Reality is closer to signal (B1, B2) with some finite rise and fall times. The analog signal (B1) is primarily a RC time constant for both rise/fall transitions. For graphic

simplicity, these are illustrated herein as ramps (B2). CMOS logic is usually designed such that T_{rise} and T_{fall} are equal. Due to semiconductor variance, they will never be perfectly matched, but are similar.

Most IC specifications define rise-fall times as the 10% to 90% time for the rise/fall period. Most digital devices have rise-fall times between 0.1ns to 5 ns. Modern FPGA devices often allow the programming of “drive strength” to adjust that time period. As will be shown, rise-fall times under 1ns can be very problematic in board designs.

Going forward, “rise time” refers to either rise or fall time. In cases where the rise and fall times are not matched, the shortest time should be used for calculations. Since digital signals are not a linear function, an easier way to analyze this is to look at rise time as it compares to propagation delay of the digital signal.

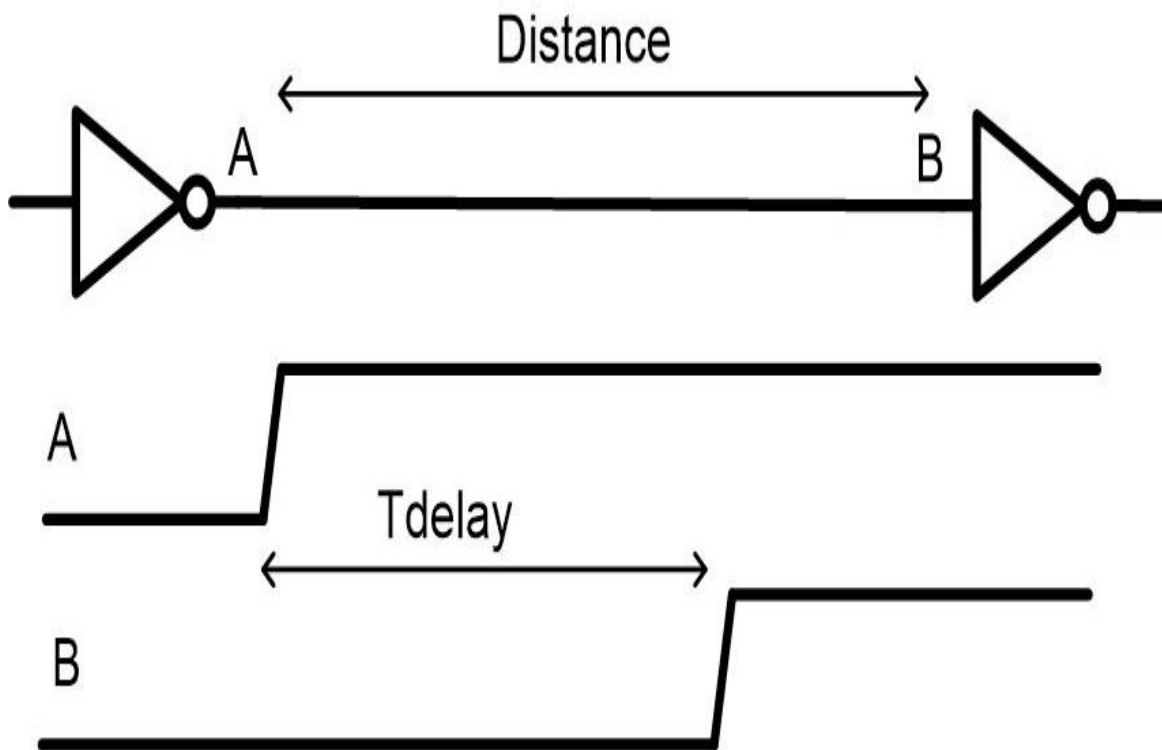


Figure 3-9. Propagation Delay Due to Connection Length

For the signal (A) sent out, it takes time for the signal to arrive at location (B). This is called the propagation delay of the signal, indicated here as

T_{delay} . For a typical FR4 circuit board, the propagation delay time is about 60-70 picoseconds per centimeter. The relationship between T_{delay} and T_{rise} is important because...

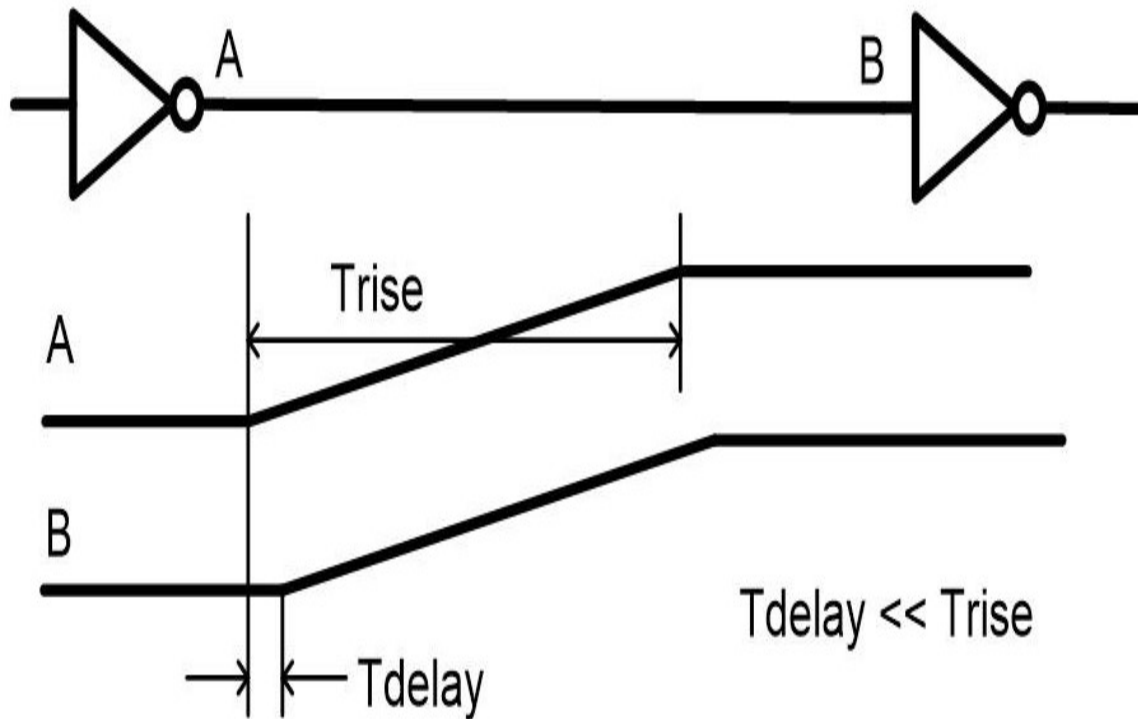


Figure 3-10. Lumped Network, Slow Rise Time or Short Path

In the above, the relationship of T_{rise} to T_{delay} is such that the propagation delay is much smaller than the rise time. Consequently, the same signal, at the same time, is seen at both A and B. Either a short propagation delay (physically short connection) or a longer rise time will do this. This situation can generally be treated as a lumped connection, A and B are the same node.

A distributed connection has a shorter rise time, or a longer propagation time.

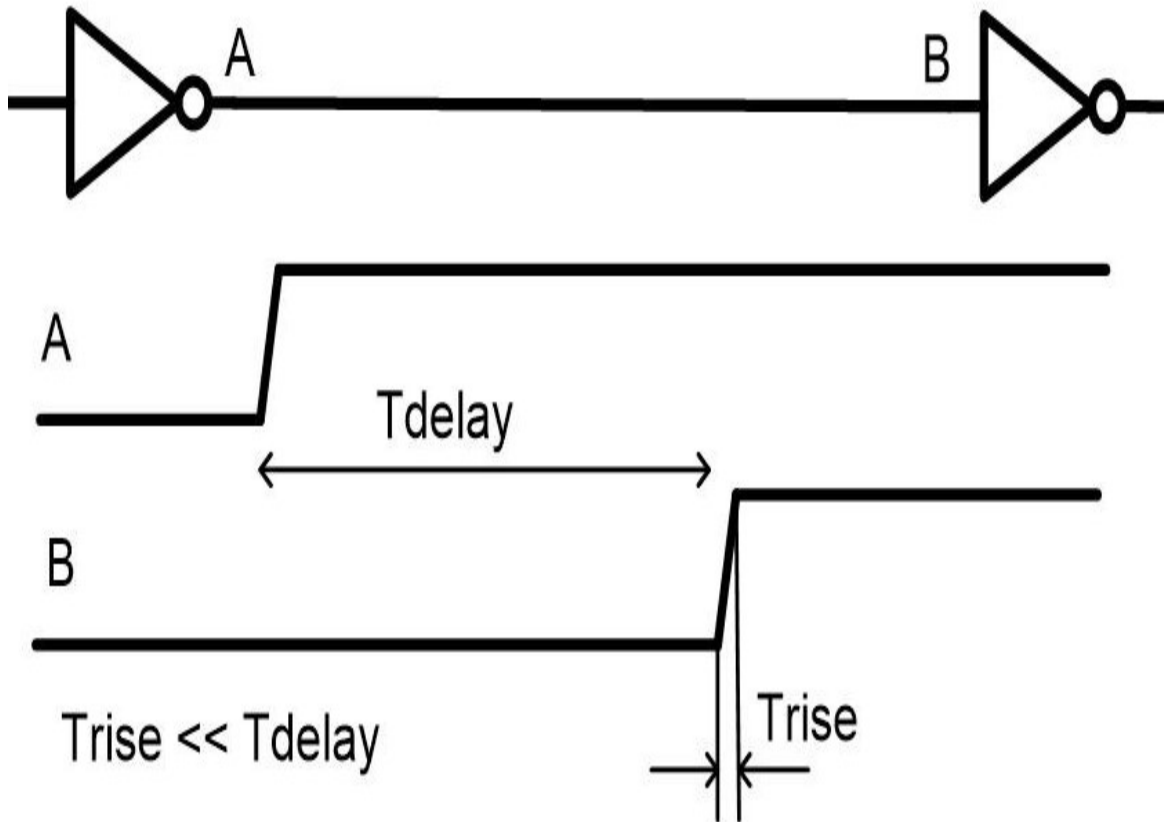


Figure 3-11. Distributed Network Fast Rise or Long Path

Here, the rise time is shorter than the propagation delay time. The voltage at point A is different from point B during logic state transitions. This needs to be dealt with as a distributed network. Also, this connection may need to be turned into a controlled impedance transmission line and need a termination resistance to avoid impedance mismatches and line reflections.

Signal length, or length of the rising edge, is the next useful tool here. Approximating a propagation delay of 60ps/cm allows estimating the physical length of a transient signal. Some common rise times for digital logic, from 0.1ns to 5ns and their signal length are shown in Figure 3-?.

Rise/Fall Time vs. Length Transient Edge

Digital Rise Time (1)	Length Transient Edge (2)	20% Length Transient Edge (3)
0.1 ns	1.7 cm	0.34 cm
0.5 ns	8.3 cm	1.7 cm
1.0 ns	17 cm	3.4 cm
2.0 ns	34 cm	6.8 cm
3.0 ns	51 cm	10 cm
4.0 ns	68 cm	14 cm
5.0 ns	85 cm	17 cm

(1) Use fastest edge for devices with unbalanced rise/fall times.

(2) Propagation time of 60ps/cm used here, typical for FR4 PCB

(3) 20% of Length of Transient Edge used, opinions differ between 16% - 25%

Figure 3-12. Rise/Fall Time and Length of Signal Propagation

The connection length must be much shorter than the signal length in order to treat a connection as a lumped network. About 20% of the signal length is the maximum connection length before the connection must be treated as a distributed network. The table shows 20% of the length of the transient

edge, which serves as a good metric to determine if a more stringent analysis is needed.

Designers rarely need a detailed analysis of a digital connection. Checking device loading vs. drive strength, using transmission lines with termination where guidelines indicate they are needed, and some common sense in signal routing generally gets it done. However, a brief look at a distributed model example is useful to better understand high frequency connections.

Distributed Model – Multi Segment for SPICE

Trace Width:	1mm
Copper Weight:	1oz
Substrate:	FR4 PCB
Inductance:	4 nH/cm
Capacitance:	1 pF/cm
Resistance:	5E-3 Ohm/cm
Signal Rise Time:	1 ns

For 10 cm long trace:

L_{total}	= 40 nH
C_{total}	= 10 pF
R_{total}	= 50 mΩ

Length of Transient Edge = $(1 \text{ ns}) / (60 \text{ ps/cm}) = \sim 17 \text{ cm}$

$$\#RLC \text{ Sections} > 10 \left[\frac{\text{Length of Connection}}{\text{Length of Transient Edge}} \right]$$

#RLC Sections = 6

For each RLC section:

L	= 6.7 nH
C	= 1.7 pF
R	= 8.3 mΩ

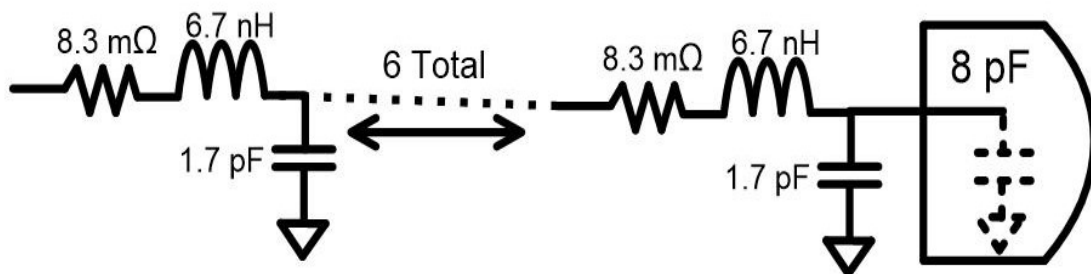
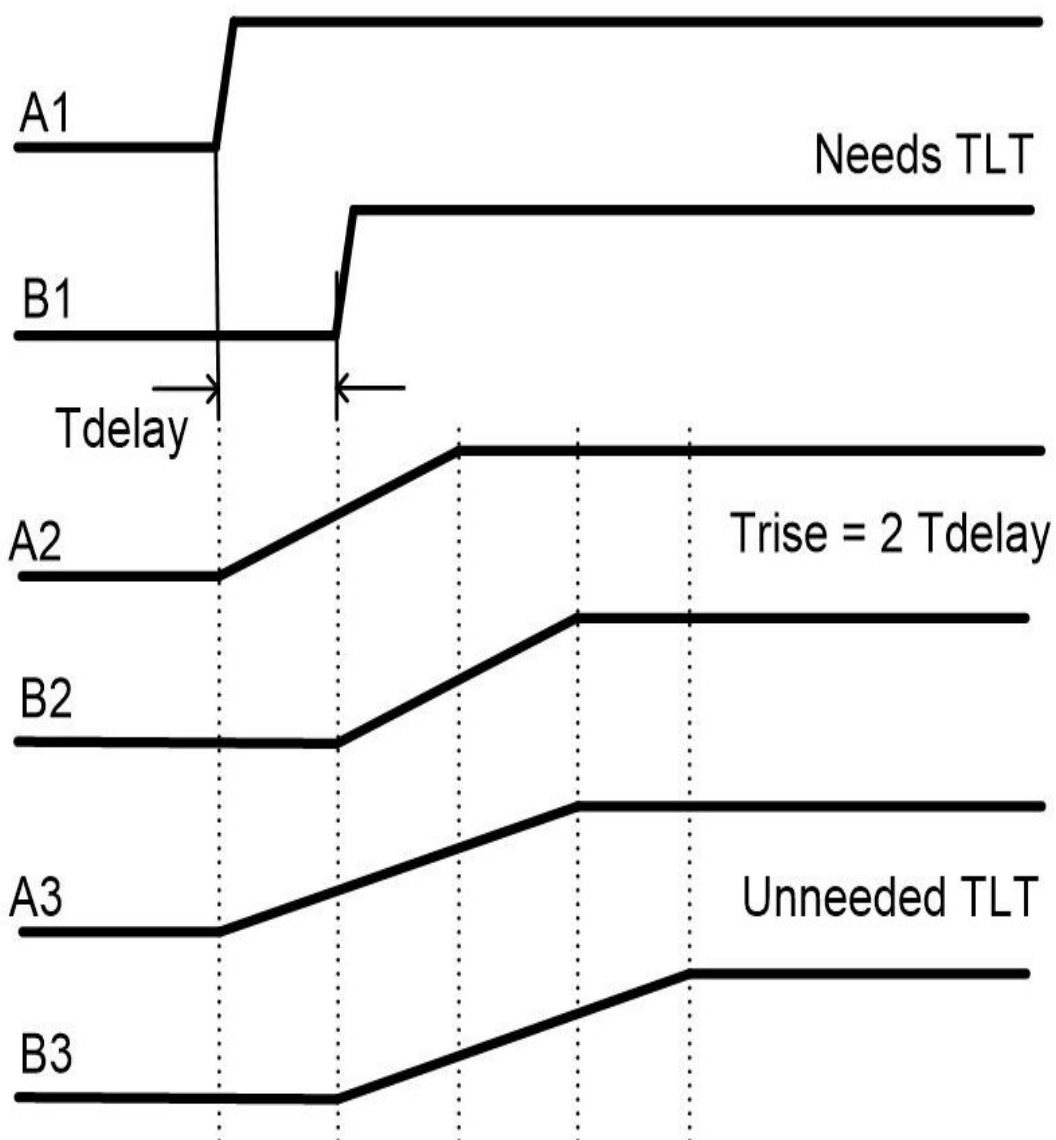
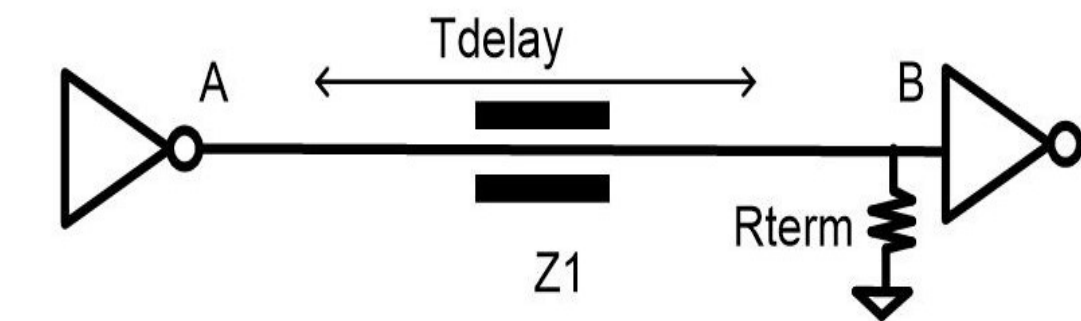


Figure 3-13. Distributed Model Example

There are multiple ways to simulate and model a distributed connection, transmission line and electromagnetic simulators are available, or, an equivalent circuit model can be created for SPICE. The equivalent circuit model shown in Figure 3-? is a 1mm wide PCB trace that is 10 cm long. Such a device would have a distributed resistance ($50\text{ m}\Omega$), inductance (40 nH), and capacitance (10 pF). Breaking this up into discrete RLC segments is an approximation of behavior. A suitable number of RLC segments can be found using the transient edge length of the digital signal. Using 10 (or more) RLC sections would be suitable for the entire transient edge length. Since a rise time of 1 ns is a signal length of 17 cm , and the connection length is 10 cm , proportionally fewer sections are suitable. A cascade of 6 RLC sections is shown, with a termination capacitance of 8 pF representing the capacitance of the receiving device. This could be simulated to provide some basic waveform information.

When to add an impedance defined transmission line with a matching termination is the next consideration.



Need TLT when: $T_{\text{delay}} > T_{\text{rise}}/2$

Figure 3-14. When to Use Transmission Line and Termination (TLT)

Using a TLT is generally done when T_{rise} is less than twice the T_{delay} . The three cases are shown, (A1, B1) needs TLT, whereas (A3, B3) does not. The (A2, B2) situation sits on the edge. A quick guideline table is useful:

Digital Rise Time (1)	Use TLT Beyond Distance
0.1 ns	0.8 cm
0.5 ns	4.2 cm
1.0 ns	8.3 cm
2.0 ns	17 cm
3.0 ns	25 cm
4.0 ns	33 cm
5.0 ns	42 cm

(1) Use fastest edge for devices with unbalanced rise/fall times.

(2) Propagation time of 60ps/cm used here, typical for FR4 PCB

Figure 3-15. Transmission Line and Termination – Rise Time vs. Distance

Use transmission lines with controlled impedance layout and resistor termination as shown in Figure 3-?. As displayed, devices with rise times

under 1ns are very restricted in a PCB environment. Most board level digital devices use rise times in the 1-5 ns region.

Fast devices require special efforts in PCB layout to avoid problems with transmission line reflections and impedance mismatches. Slower rise time devices are easier to work with. Modern FPGA/CPLD devices often have a feature to select the “drive strength” of an output pin. Using weaker drive strength gives slower rise times and less problems. If a slower device is not viable, or adjusting drive strength is not an available tool, there are other options.

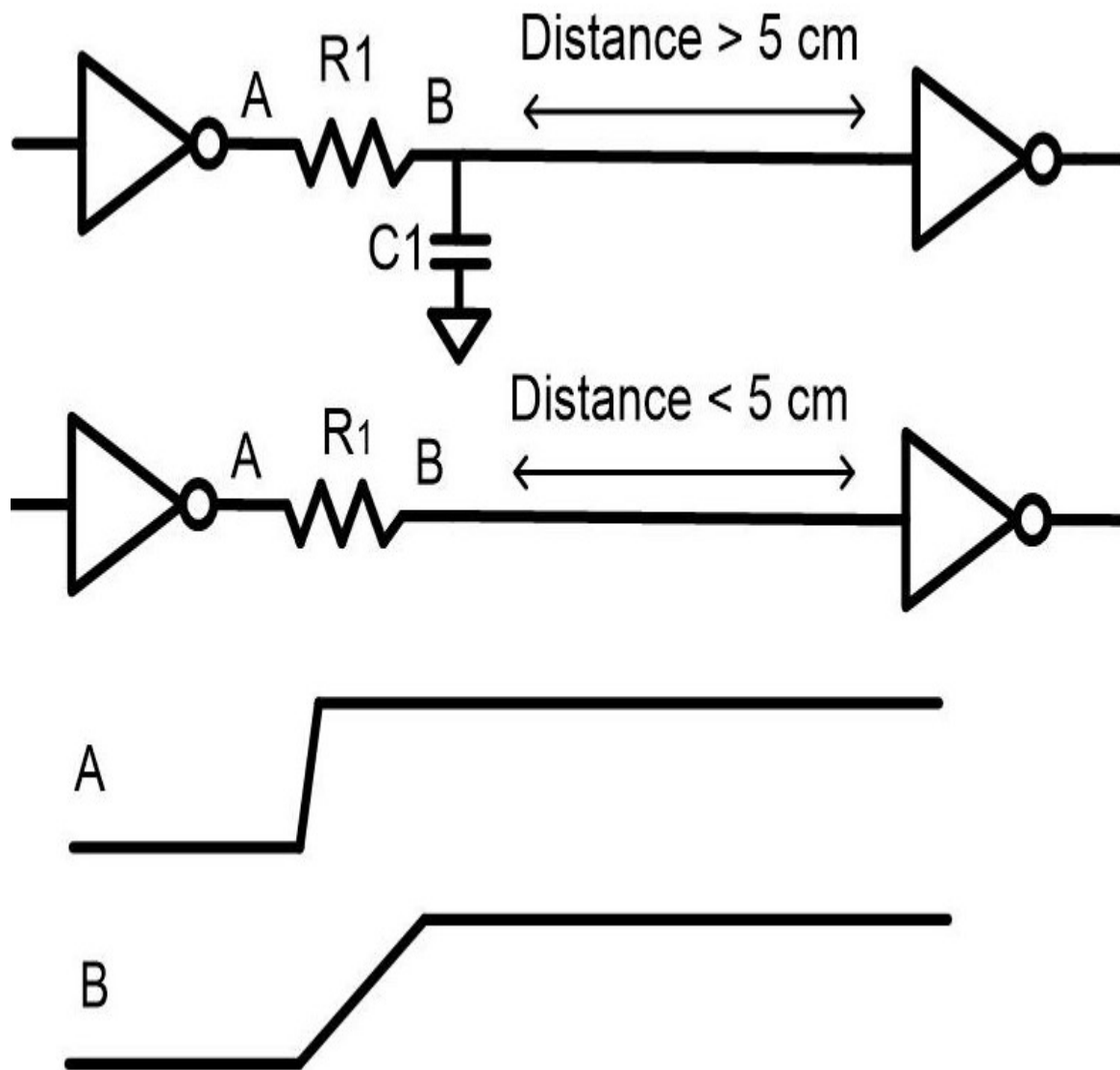


Figure 3-16. Increasing Rise Time to Allow Longer Path Transmission

Inserting resistance at the driver end of the connection can slow the rise time enough to create a longer connection. Short connections (<5 cm) can use the parasitic capacitance of the PCB and the IC being driven. Longer connections (>5 cm) need a capacitor closely placed with the resistor, to avoid fast transients on a long inductive connection. This can be useful in certain situations where a longer distance connection is needed and other methods are not readily implemented. This will skew signal timing slightly, but small amounts of resistance (10-100 ohms typically) can change the rise time enough to fix a problematic signal without redesigning the circuit.

Remember that ground referenced digital signals have limitations, even when run through properly terminated transmission lines. LVDS methods through differential terminated transmission lines are preferred when speed, distance, and noise immunity are important.

Clock Distribution

Low frequency ($<1\text{MHz}$) clock distribution should be straightforward. As frequency goes up, connection distance increases, and more destinations are added, the signal integrity, timing jitter, and time skews between destinations need to be analyzed. Using a comparison of two cases with commonly used clock frequencies illustrates the following issues:

A Real Time Clock (RTC) runs at 32.768 KHz . This frequency is suitable for crystal fabrication and easily divided to create a 1 second clock. Using a RTC is a common strategy for many controller systems. The RTC period is $30.5\text{ }\mu\text{S}$.

The PCIe interface standard uses a 100 MHz clock that is distributed as part of the PCIe bus. The PCIe Clock (PCIC) period is 10 nS .

Clock Buffering

A 5 nS gate propagation delay is typical for a clock buffer. For the RTC, that is 0.016% of the period, which is insignificant and can be ignored. For PCIC, the same buffer is 50% of the period and significantly affects timing.

Propagation Delay

A 50 cm long PCB connection has a propagation delay of about 3 nS (using 60ps/cm). The RTC gets a clock skew of 0.01% of the period and the PCIC skews 30% of the clock period.

Rise/Fall Time

As an arbitrary number, allocate 5% of the clock period to each rise/fall cycle. This means the RTC needs rise/fall times less than $1.5\text{ }\mu\text{S}$, and the

PCIC needs 0.5 nS. The PCIe system needs a transmission line beyond 5 cm. The RTC can have slower rise times and avoid a TLT for many situations. (Side Note: Many logic devices exhibit metastability with slow transition edges, so device specifications for slowest transition edges or using devices with hysteresis needs to be investigated.)

Both propagation delay and gate delay can seriously affect accuracy, or can be ignored as inconsequential, depending on clock frequency. The path to take can be determined by looking at the clock frequency and determining performance similar to the above.

Too much capacitive loading can degenerate signals. Logic vendors use the concept of output driver “fan out” and input device “standard loads” to simplify the problem for designers. For CMOS devices “standard loads” are a simple method of adding up the capacitance attached to a signal.

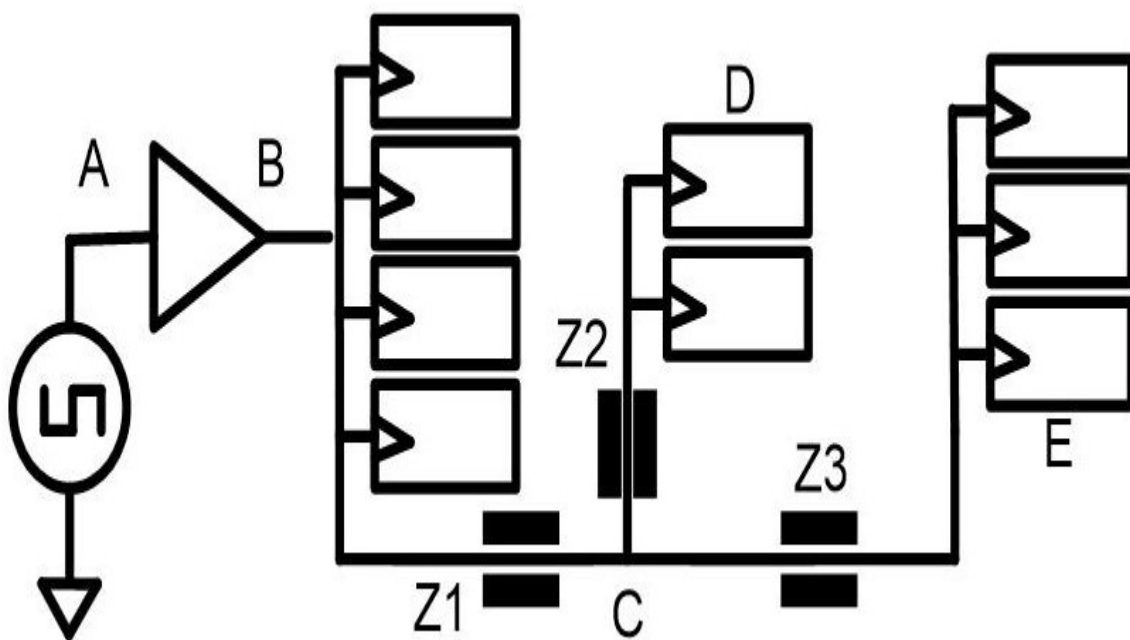


Figure 3-17. Skewed clock trees multi path edge degeneration

In Figure 3-?, a clock at (A) passes through a single clock buffer and (B) is distributed to 9 locations with long connections (Z1, Z2, Z3) and no path

organization. The first problem here is capacitive loading presented by 9 different devices. Check the capability vs. loading of the clock buffer drive and the loads due to the 9 devices and interconnect. If the setup has low clock rates and slow rise times, this may be sufficient, but there is room for improvement.

The second problem here is the existence of a long stub (Due C to D and Z3) in between two other long signal paths (Z1, Z3). Connections branching off other connections are often a problem with impedance mismatches and line reflections. Single point to point connections are highly preferred. Cleaning up signal paths can produce more predictable behavior:

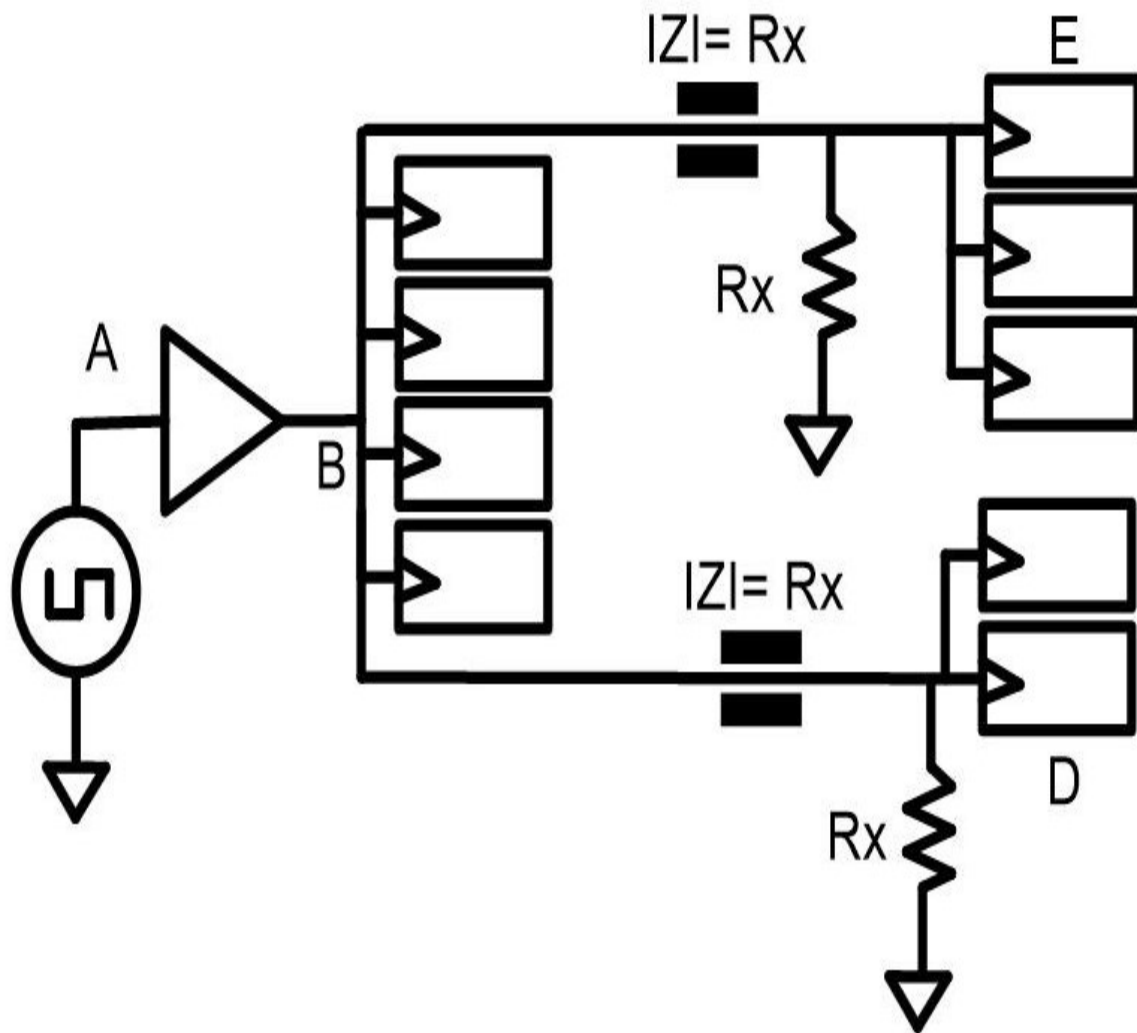


Figure 3-18. Spider Network of Clock and driving multiple lines

This structure has the same tightly grouped loads at B, D, E. Instead of a branched connection there are two separated transmission lines (B to E and B to D) and no stubs. The two paths have been replaced with a TLT to minimize mismatch reflections. Several items still need to be either checked or resolved:

- Examine the total driver loading with 9 logic loads, transmission lines, and DC loading from termination resistors.
- Transmission lines (B to E) and (B to D) share a common node, and can interact and cause signal integrity issues.
- Timing skew errors between signals at B, D, E, need to be checked.
- Rise/fall characteristics at B, D, E will be somewhat different due to different impedance and loading, may or may not be significant.

Minimizing these issues by using a multi output clock driver is often a solution.

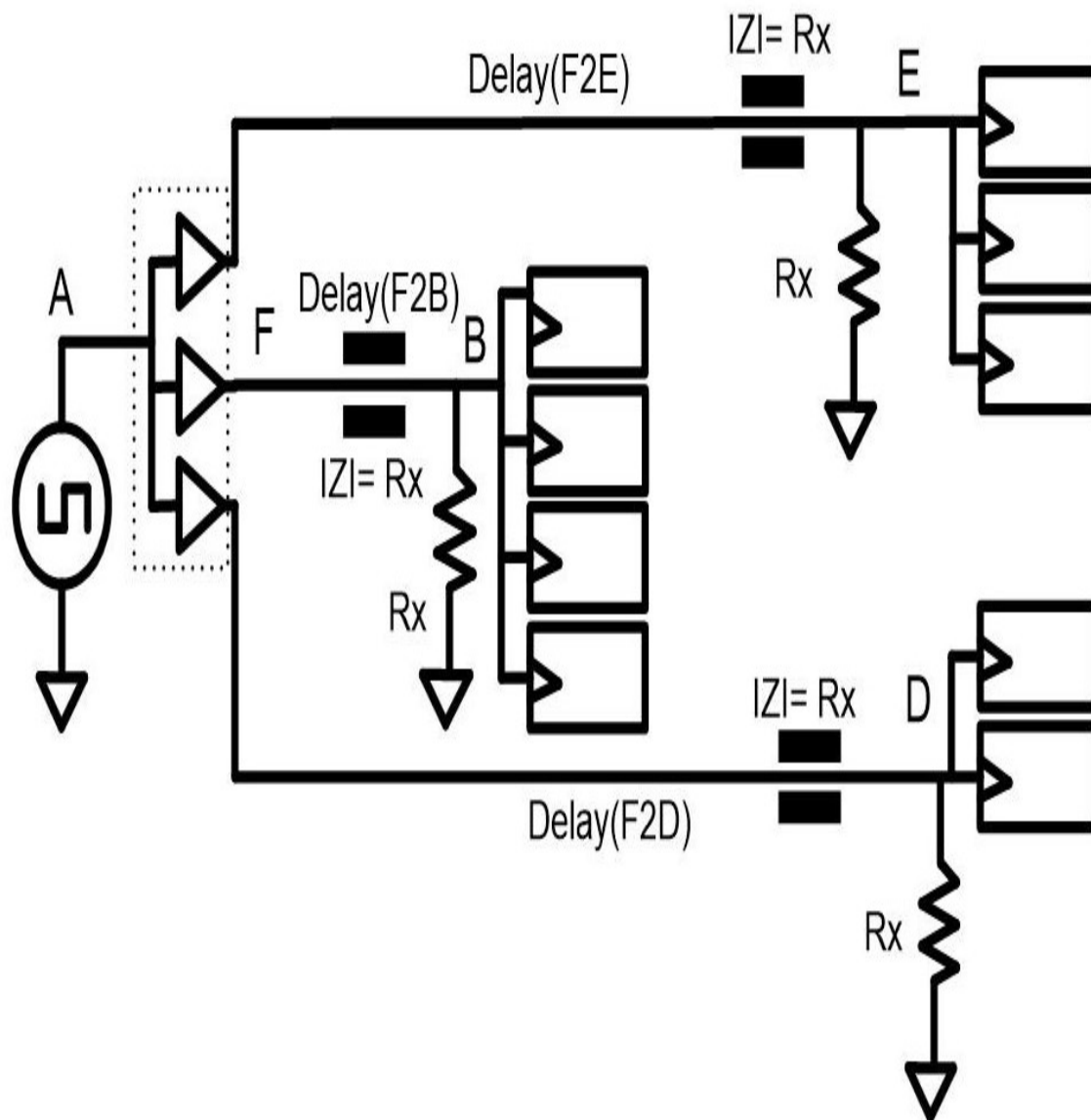


Figure 3-19. Clock Distribution With Buffer Path/Load Balancing

In Figure 3-?, three separate drivers with TLT are used. Interactions between transmission lines are eliminated and driver loading is reduced. Signals at B, D, E, should have similar rise-fall characteristics. Clock buffers are available with timing skew errors under 100pS (ground referenced) and under 10ps (LVDS) variants. The importance of timing skew errors due to physical length mismatch of F2E, F2B, and F2D needs to be checked.

Do F2E, F2B, and F2D need to be length matched in order to match propagation delay time? Transform the length differences into time differences, and compare to the clock time period of the clock to determine if its physical path balancing is a significant time skew component.



Figure 3-20. Serpentine path balancing by the numbers

If length differences are an important part of the clock period, serpentine interconnects in the shorter paths can reduce time skew. Serpentine paths take up significant space, and frequently are unnecessary. Important questions here are:

- Is the path length difference a significant portion of the clock period?
- Is a time skew tolerable between separate parts of the system?

Older designs, with many logic gates, frequently needed careful clock path balancing. Modern systems may have a single FPGA as the destination and careful path balancing is not needed.

Time skew can also be created with:

- series resistance creates delay through use of RC time constants
- gate delays in series

Gate delay can change due to semiconductor process variance and needs to be carefully examined over specified variances. Introduction of limited amounts of series resistance produces a more consistent result but time shifting the clock should be kept under 10% of the clock period to allow sufficient time for the RC time constant to settle out.

As a general rule, avoid introducing intentional signal delays unless there is a compelling need for it. In the single FPGA era, most of the time, it is unnecessary.

If shifting a signal to an earlier time is needed, zero delay buffers are also available.

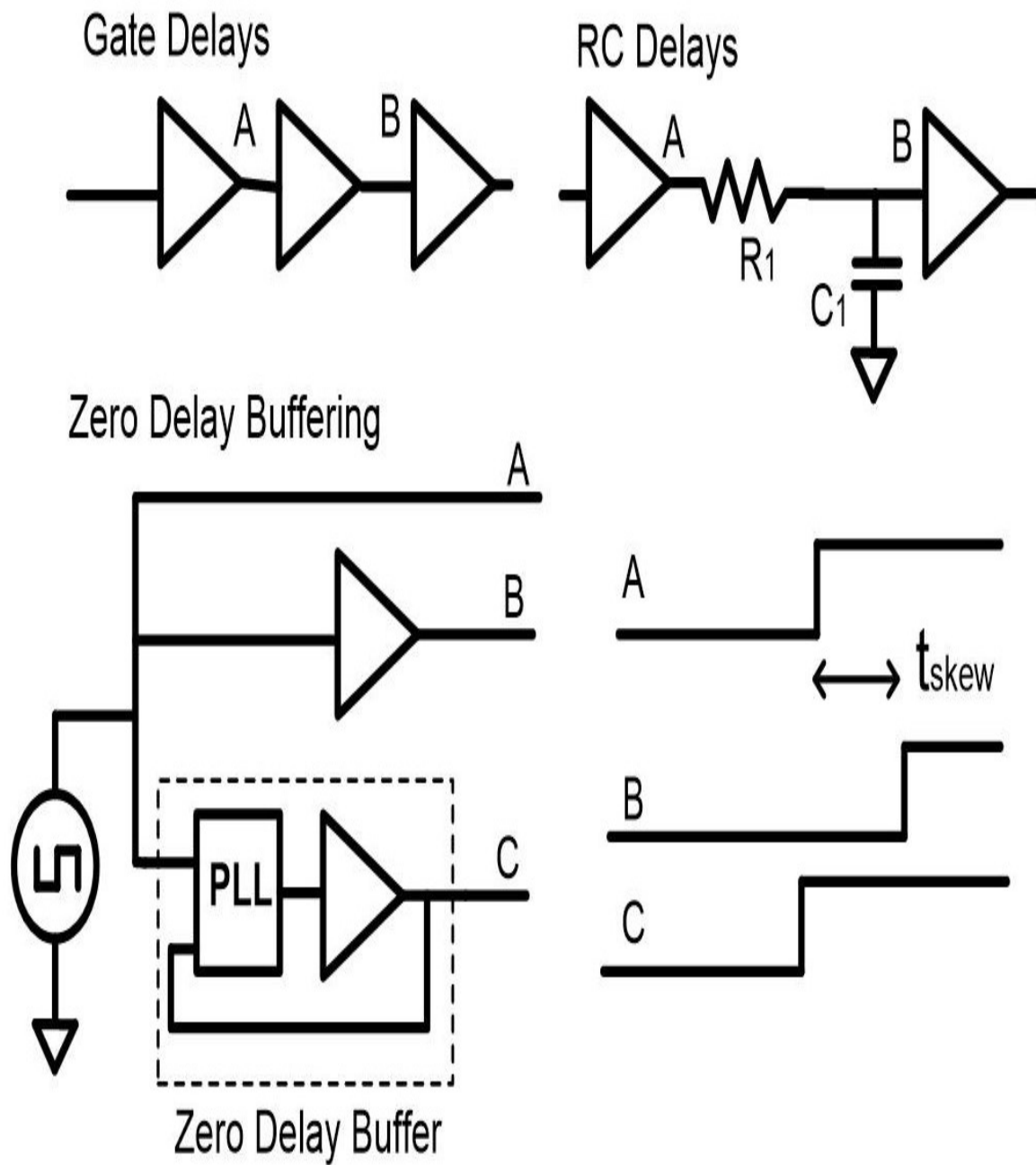


Figure 3-21. Skewing the clock delay methods and zero delay buffering

The zero delay buffer output occurs at the same time as the input. To do this, a Phase Locked Loop (PLL) is used internally to regenerate a clock output with zero time skew to the input signal.

High frequency clocks within noisy environments should consider using a LVDS method.

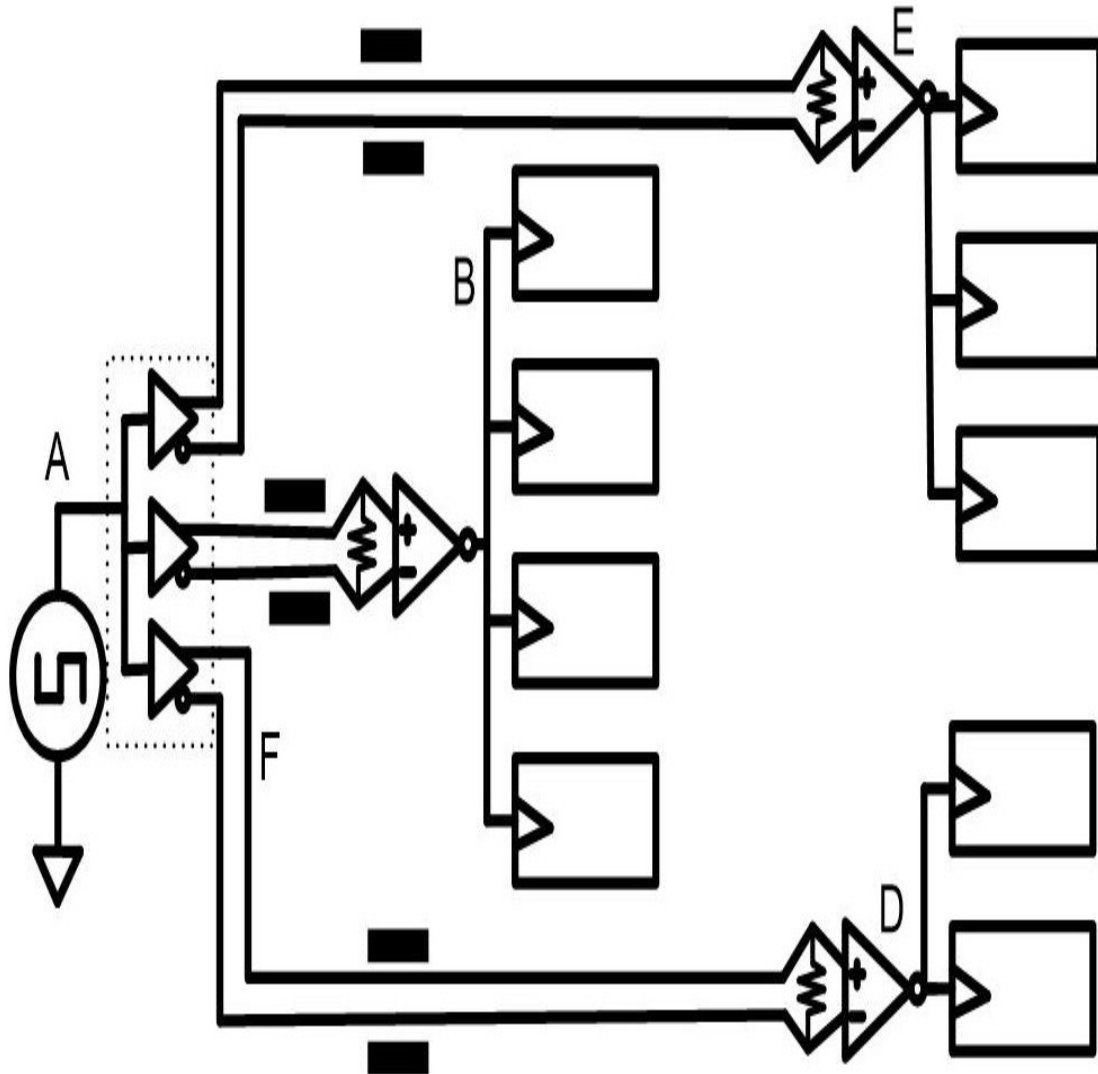


Figure 3-22. Clock Distribution with LVDS Transmission

An LVDS clock transfer using differential transmission lines and impedance matched differential termination works well for high frequencies and noisy applications. PCIe uses this method to distribute the 100 MHz reference clock required by the standard.

Digital Communication - Parallel vs. Serial Ports

Data can be transferred between devices in a parallel or serial manner.

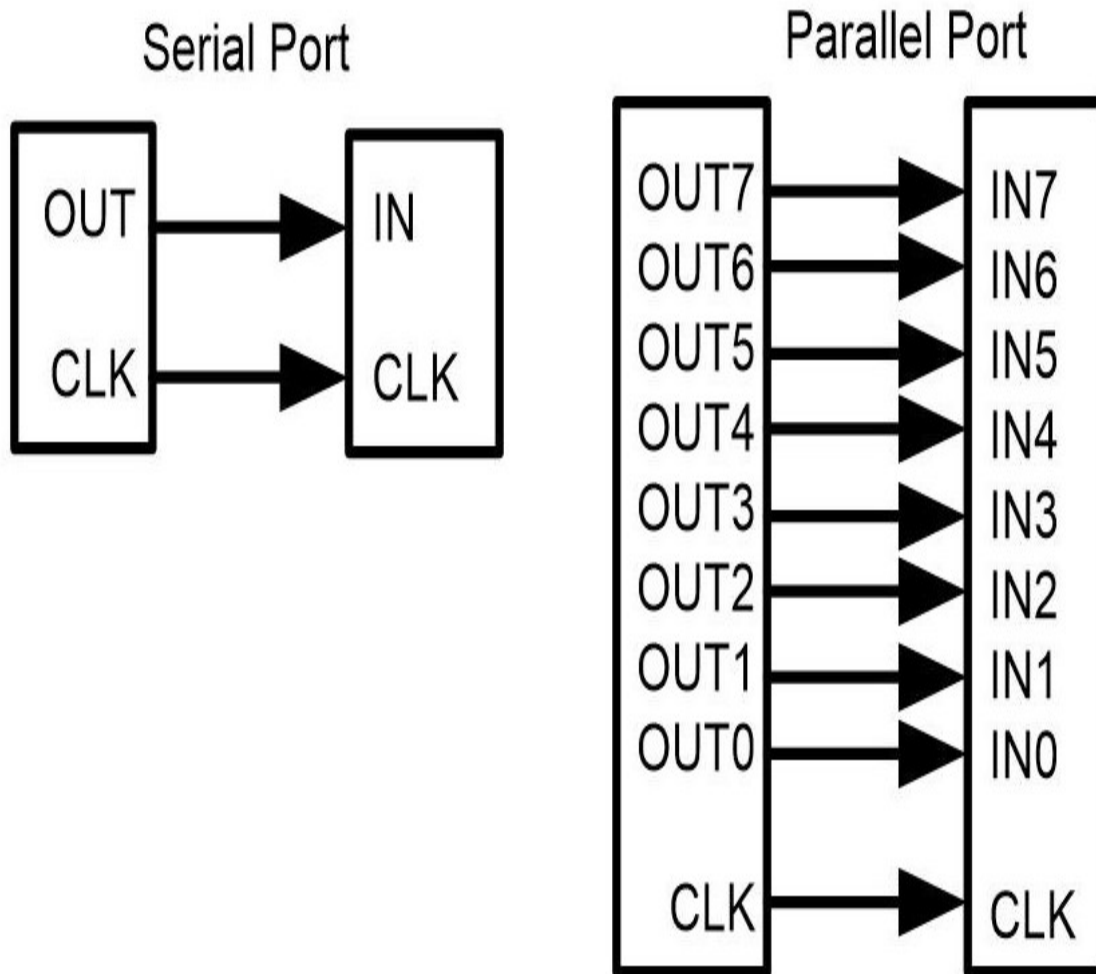


Figure 3-23. Parallel vs. Serial Communication

Parallel communication requires multiple pins and connections between devices. A large amount of area inside both IC's is used to provide the external signals. Due to silicon cost and complexity of interconnects, parallel ports are obsolete. Serial Ports pass data sequentially between devices with a reduced pin count. Modern systems use serial port methods, and are the focus here.

Clocking Methods for Serial Ports

In addition to passing data, some method of providing clock synchronization must be used. The four methods applied are: starting edge synchronization, parallel path clock, self clocking data, and embedded clocking.

Clocking Methods – Starting Edge Synchronize

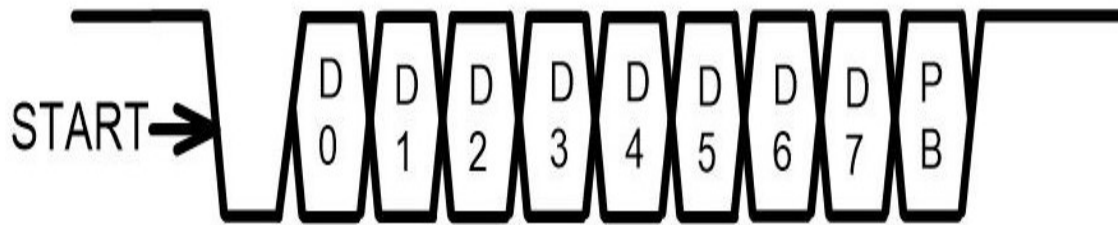


Figure 3-24. Starting Edge Synchronization

Starting edge synchronization requires that the transmitter (TX) and receiver (RX) have a predefined data rate, and access to a reference clock. In this example, when TX goes low, the RX recognizes that as the start edge of the signal and uses its reference clock to sample the expected middle of the data window for a short series of data bits. Clock phase errors get progressively worse after synchronization so this only works for slow and short data fields. Error prone and originally used in RS-232 and UART devices, it is not encourage for new designs.

Clocking Methods – Parallel Clock

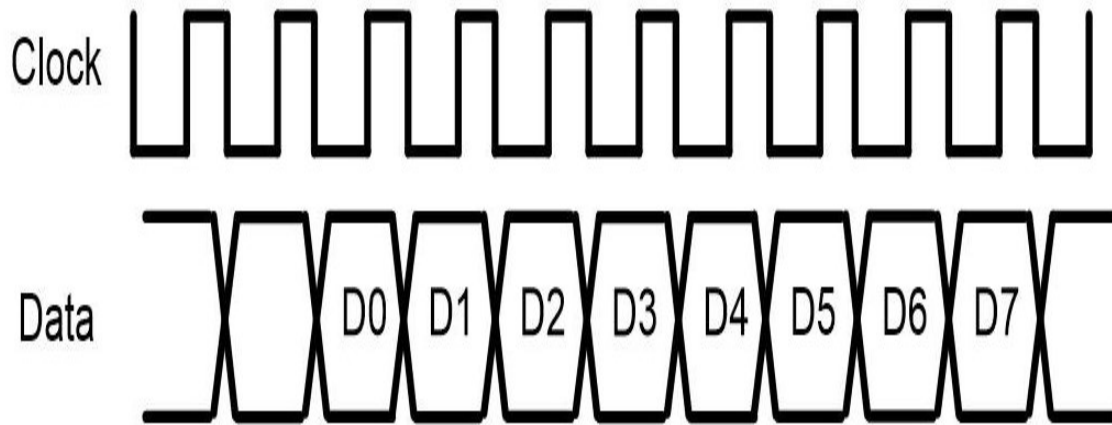


Figure 3-25. Parallel Clock

A parallel clock is time synchronized to the data. Clock to data needs to be properly phased for sufficient setup and hold times. The limitations of parallel clock systems are the need for a second connection and maintaining a good clock/data phase relationship. High speed data systems can have problems maintaining this relationship and prefer to use embedded clocking methods.

Clocking Methods – Manchester Code Self Clocking

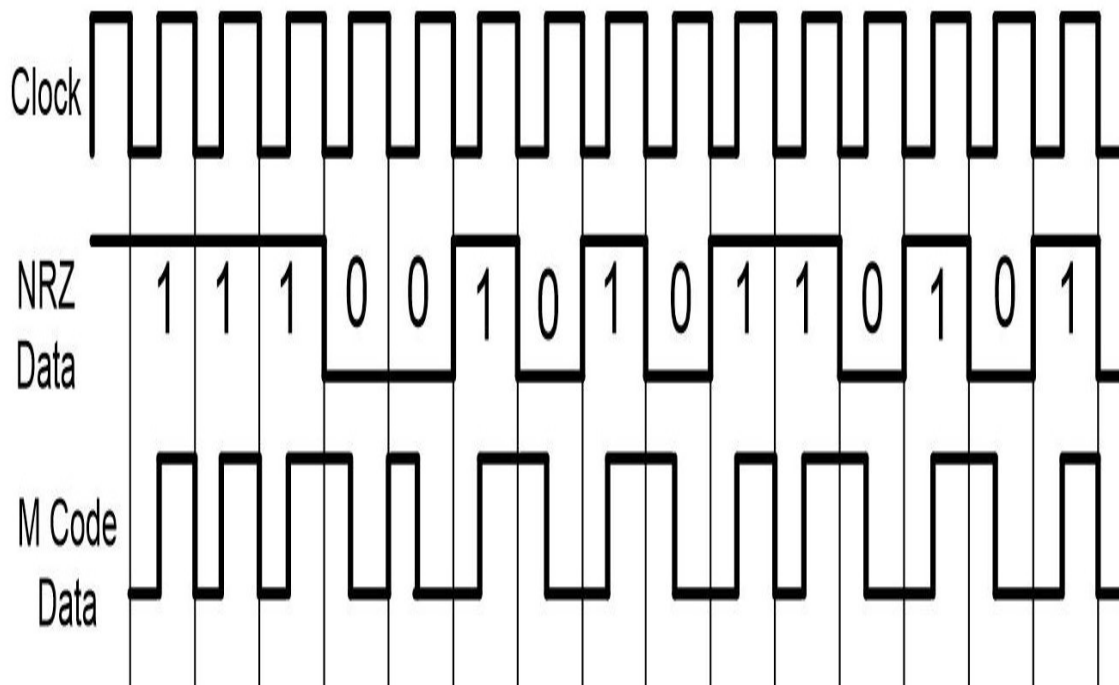


Figure 3-26. Self Clocking

A self clocking data stream is commonly created with a Manchester code. NRZ data is transformed so the data is represented as rising or falling edges. In the positive edge version of the Manchester code, a binary “1” becomes a rising edge, and a binary “0” is a falling edge. (The code set is inverted for the negative edge version) This set of transitions can be decoded and synchronized at the receiver without a reference clock. Manchester coding doubles the bandwidth of the signal, which can be problematic at high data rates.

Clocking Methods – Embedded Clock, Run Length Limited Codes

Embedded clock systems create a clock based on the data stream. This is called “clock recovery” and is accomplished using a phase locked loop (PLL). The PLL generates a clock that is in alignment with the transient edges of the incoming data. The received data needs frequent transient edges to keep the PLL properly aligned. This can be a problem with NRZ

data, which can have long strings of the same state and lacks transient edges for timing alignment.

Avoiding a long string of ones or zeros requires encoding the NRZ transmission data. NRZ data, is converted to “Run Length Limited” (RLL) data which guarantees a min/max limit between transient edges. This allows successful, phase aligned, clock creation.

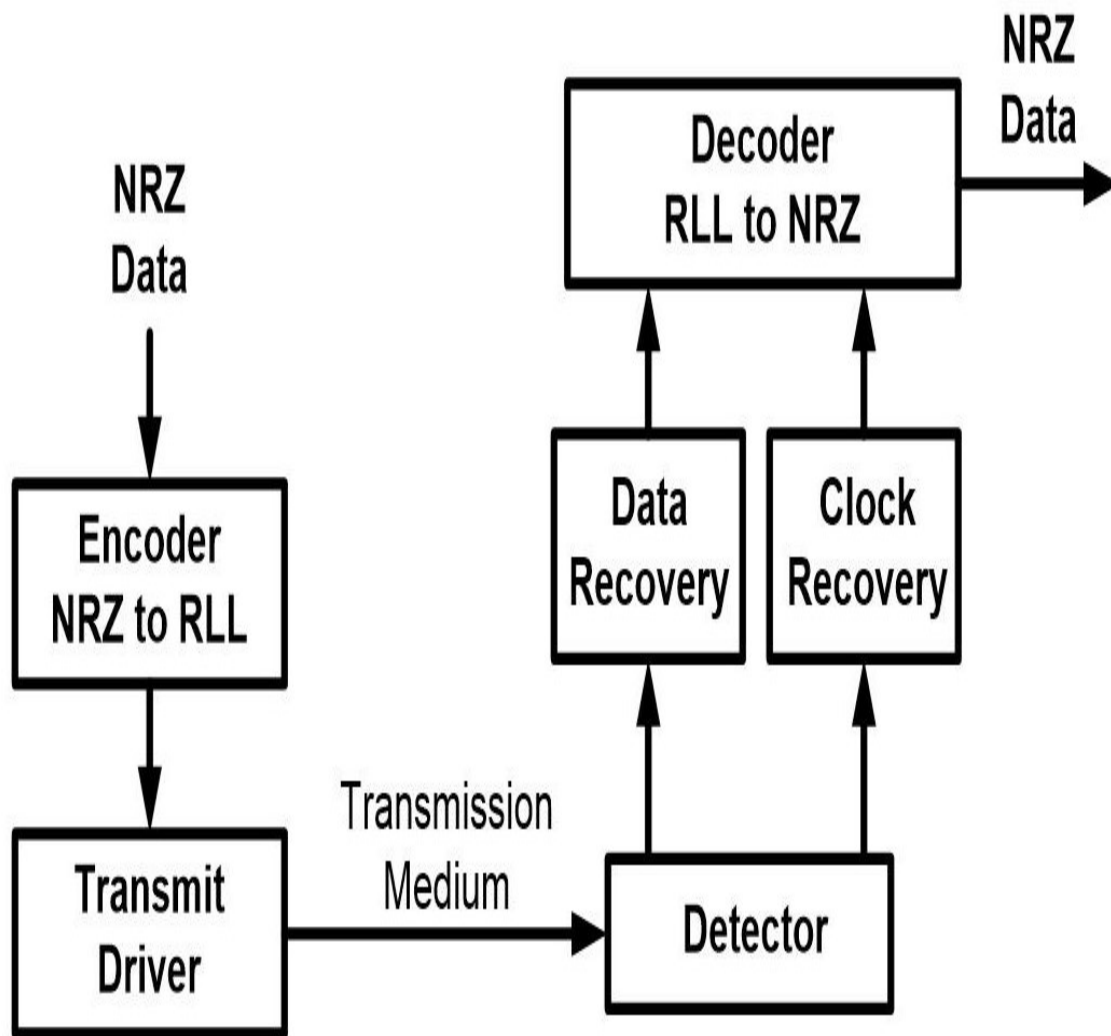


Figure 3-27. Embedded Clock System

Figure 3-? illustrates a typical system. The RLL data controls a transmit driver, the data passes through a transmission medium, and is detected asynchronously at the receiver. This data stream is used to recover a clock

and provide synchronous data to a decoder which converts the data from RLL to NRZ. Most high speed data communication systems use something similar to this.

All of these clocking methods are commonly used with serial interfaces.

Digital Communication - Features and Definitions

Before examining the communication methods used commercially, understanding some commonly used criteria and terminology is useful:

Data Rate

Sometimes called data transfer rate. How much data can you get through a connection per unit of time? This can be measured in bits per second, bytes per second, baud, or symbols per second. Vendor specified data rates may not include supplemental time used for support communication protocols or other overhead processes.

Distance

Any communication method is distance limited. Often, data rates need to decrease as distance increases.

Addressing

How many devices allowed on a connection is limited by digital addressing, capacitive loading, or other system specifics.

Error Checking

Data integrity is an important part of communication. Data errors between two closely placed chips, on the same PCB, is generally not a problem. But, as distance increases, or data rates go up, or noise is introduced, then error rates go up and checking integrity is needed.

Protocol, or, Protocol Structure

The pre-defined procedure that devices use to communicate. Protocol features can be very different depending on the application.

Common features include definition of: data sequence and size, encoding methods, error checking methods, acknowledgement and handshaking procedures, addressing structure, requests for start/stop/repeat, and others.

Data Packet

The block of digital information that is communicated, as defined by the protocol. A generic data packet is shown in Figure 3-?.

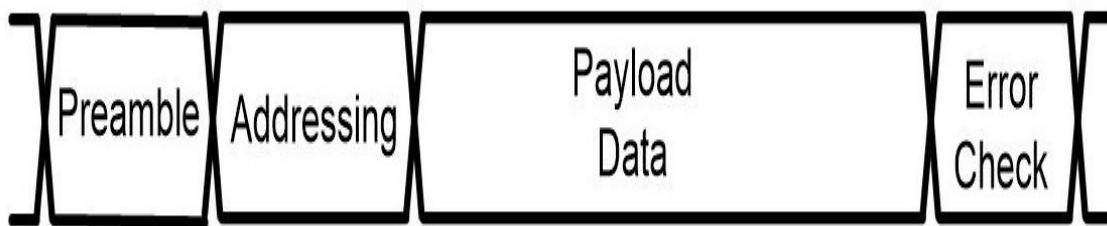


Figure 3-28. Typical Data Packet Content

Common components of a data packet include:

Preamble

Used to prepare the communications path for what is to follow. Receivers use the preamble to make signal processing adjustments, and synchronize a clock recovery PLL to the incoming data.

Addressing

Includes a destination address and sometimes includes a source address. Depending on the protocol, this can be a hardware defined register address, an IP address or a MAC address. Addressing is needed to define the destination, and in some protocols, the original source of the data.

Payload Data

Also known as “data” or “data field” and contains the actual information being conveyed. Depending on protocol, this can be fixed or variable in size.

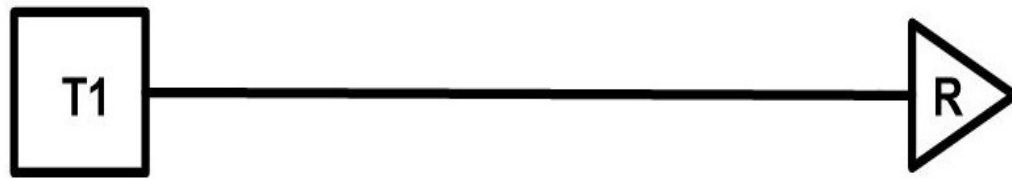
Error Check

Information exists to quickly determine if the data field was delivered correctly. Error checking is commonly done with a CRC (Cyclic Redundancy Check) code.

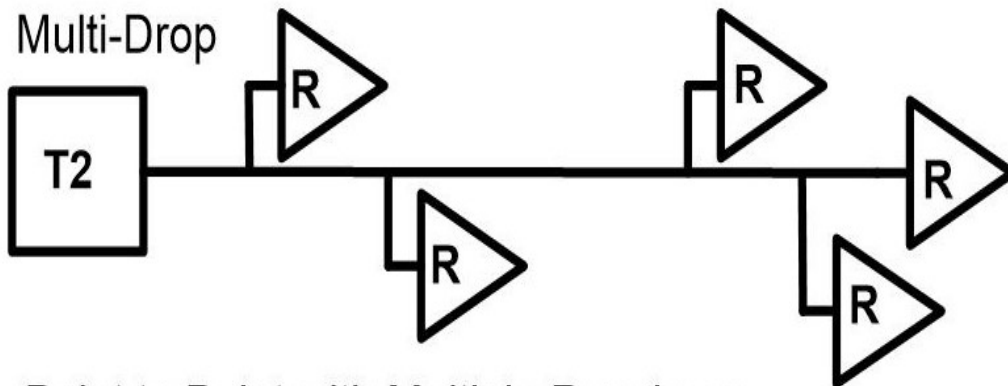
Connection Configuration

can have several forms. A Point to Point (P2P) configuration is two devices communicating, with one on each end of the connection. P2P connections are high frequency and transmission line friendly. A Multi-Drop system is a single driver with multiple receivers. A Multi-Point system has multiple receivers and multiple drivers.

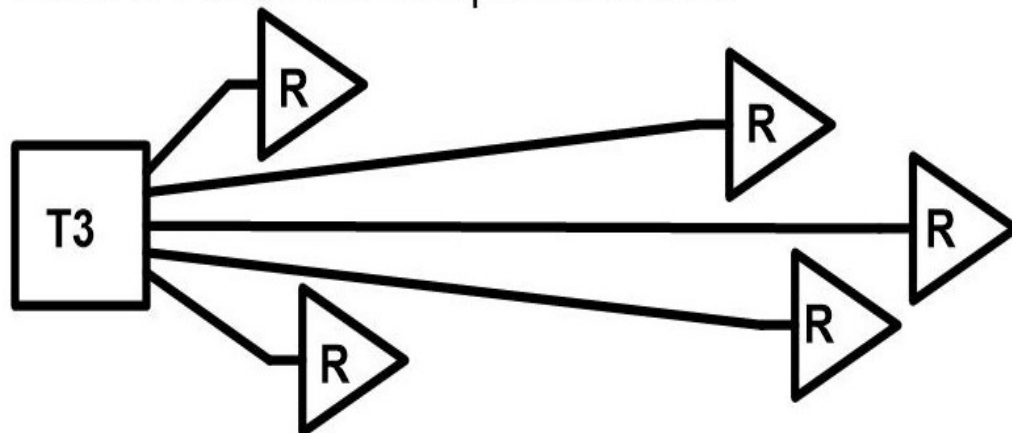
Point to Point



Multi-Drop



Point to Point with Multiple Receivers



Multi-Point

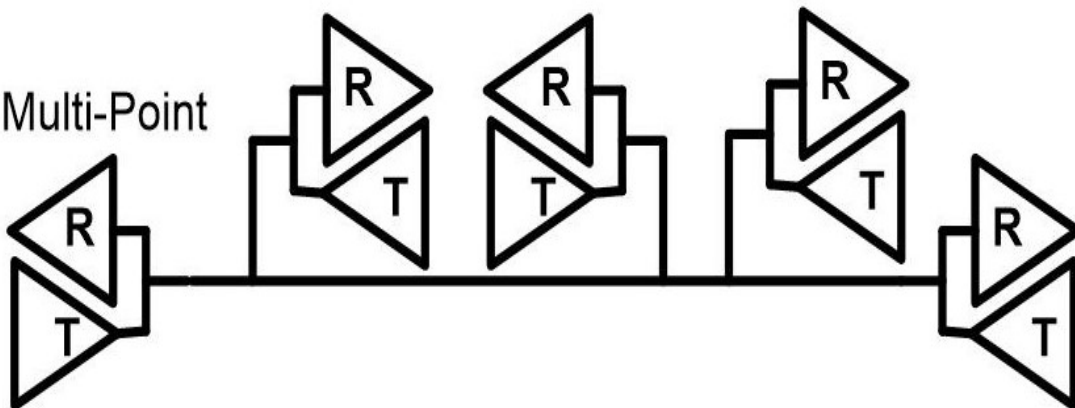


Figure 3-29. Multi-Point vs. Point to Point (P2P) Interconnect

A system can have multiple receivers while still having a P2P configuration. Separate drivers for each connection are utilized to separate connections from each other.

Hot Plugging

Refers to plugging and unplugging a device from a powered up system. Many devices can't do this without damaging the electronics. Properly designed hot plugging can insert or remove without damage, not disturb normal host functionality, and can seamlessly self configure.

Physical Medium

The connection between transmitter and receiver has multiple methods including: PCB traces, discrete wires, twisted pair wires, coaxial cable, optical fiber, infrared light, magnetic coupling, and wireless radio.

Communication Channel Capability

Telecommunication terminology still gets used to define the capability of the data path.

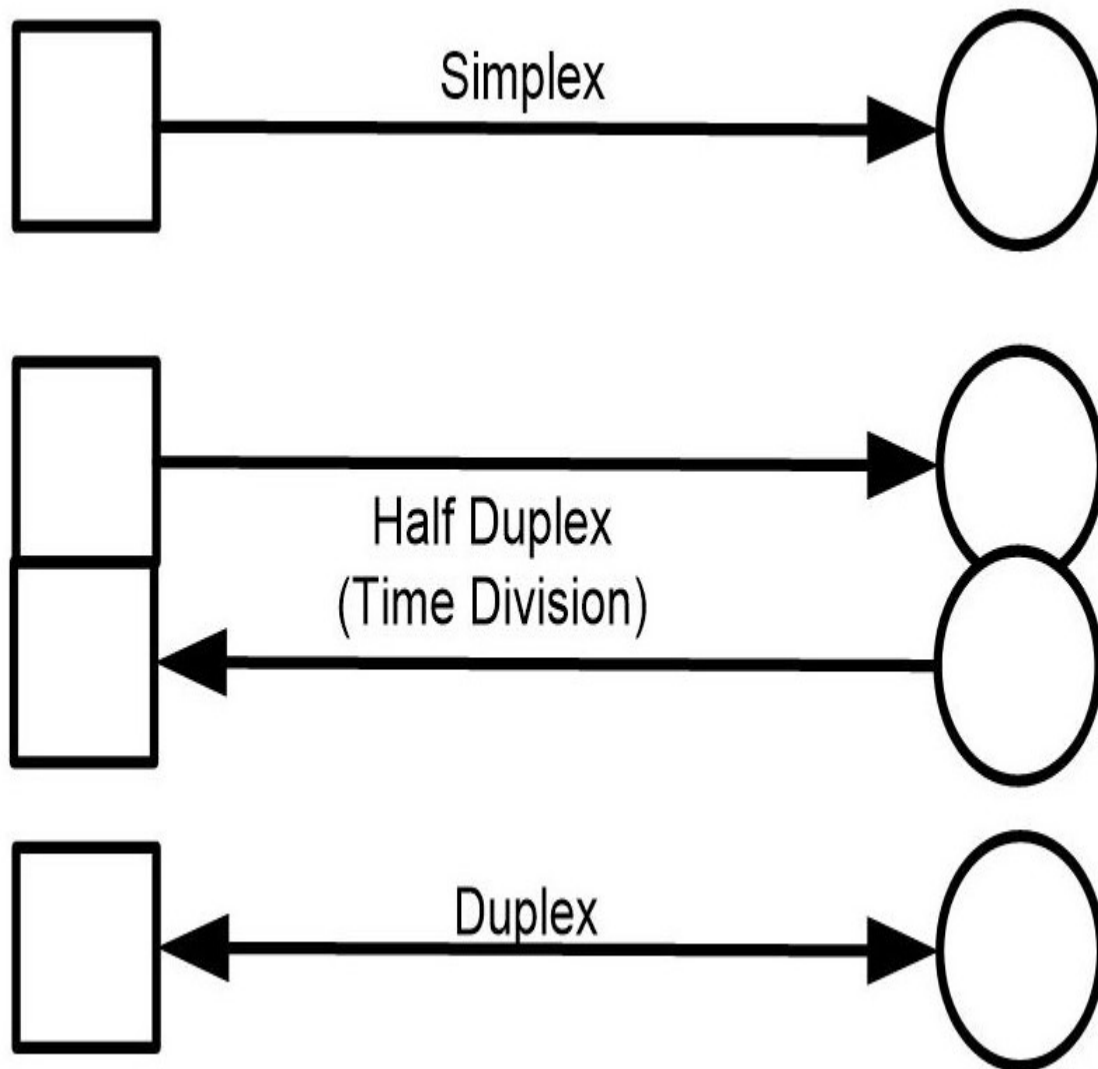


Figure 3-30. Simplex, Duplex, Full Duplex, Half Duplex Communication

Simplex systems communicate only in one direction. Loading data into a register is an example of a simplex process if there is no way to verify or read back the data.

Half Duplex systems can communicate in both directions but not at the same time. Interleaving in time is a common strategy to implement a half duplex system.

Duplex or Full Duplex systems can communicate in both directions at the same time. Depending on the device, the two way communication can occur on the same path/channel/wires or a second path can be used to

communicate in the other direction. Two simplex systems can be used together to create a duplex system.

Serial Data - Shared Ground, Low Speed

Although ground referenced logic can have signal integrity issues, it is the backbone of most digital systems. A low impedance common ground is a must have for such systems to function reliably. Some digital interfaces offer no detection and correction for errors, so they need to operate in a shared ground environment with high data integrity.

Universal Asynchronous Receiver Transmitter (UART)

The UART structure uses start edge synchronization, and is a simplified variant of the RS-232 interface.

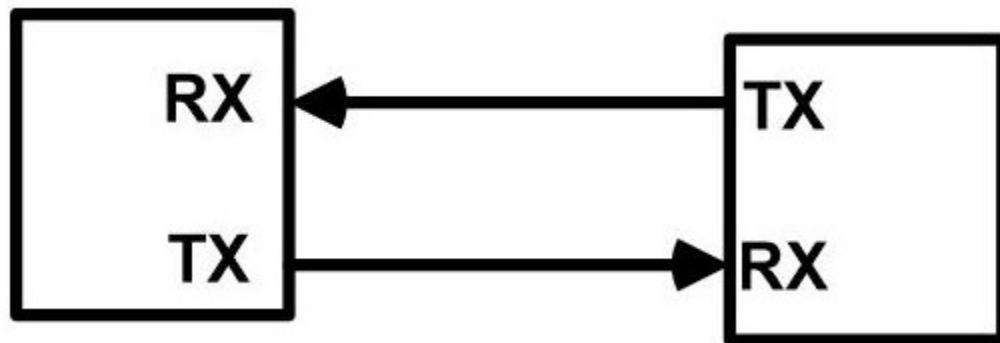


Figure 3-31. UART Connections

The UART has a receiver to transmitter connection for both directions and nothing else. Device clocking is performed internally using two independent time references where the clock time period (referred to as a baud rate) has been programmed into both devices.

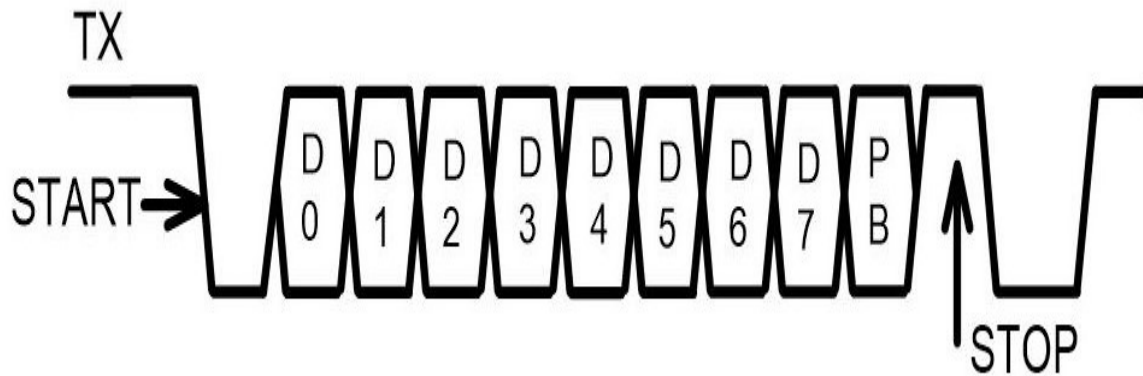


Figure 3-32. UART Signaling Strategy

Data transfer is started by the falling edge of either TX output. This falling edge (START) provides a zero phase reference for receiver synchronization. By knowing the baud rate of the expected data, the RX internal clock can time sample data in the expected middle of each bit for 9 bits. Since independent TX and RX clocks will differ in frequency, the sample point timing will drift off center after start synchronization, which limits both maximum data rate and length of data field. This method works for short runs of data only. A parity bit (PB) is added to the end of the data to determine if an odd number of errors occurred. An even number of errors would not be detected.

a common PCB. Many semiconductor vendors provide an I2C interface for their products to do setup and control of internal registers. The System Management Bus (SMB, or SMBus), developed by Intel is a variant of the I2C interface. In general, SMB and I2C devices are compatible on the same connections.

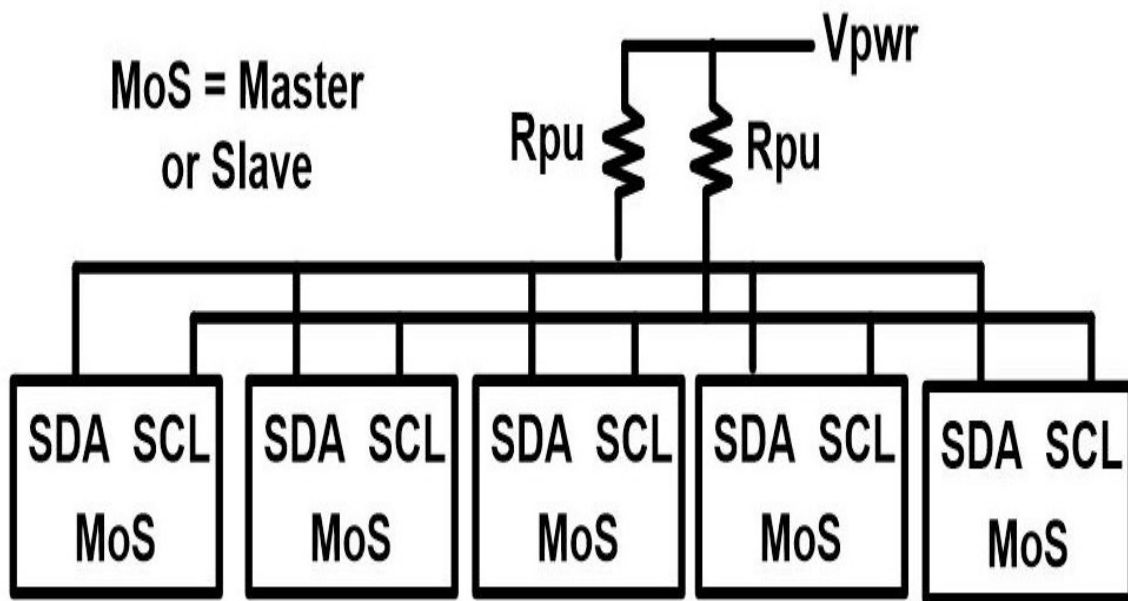


Figure 3-34. I2C Connections

I2C consists of two signals, a clock (SCL) and serial data (SDA). Multiple devices can be connected to the SDA/SCL signals, and a single set of pull-up resistors (R_{pu}) are also connected. The connection pins (SDA, SCL) of any attached chip can pull down the signal with the resistors providing a passive pull up. Any device can initiate communication and becomes the Master, controlling communication with a Slave device.

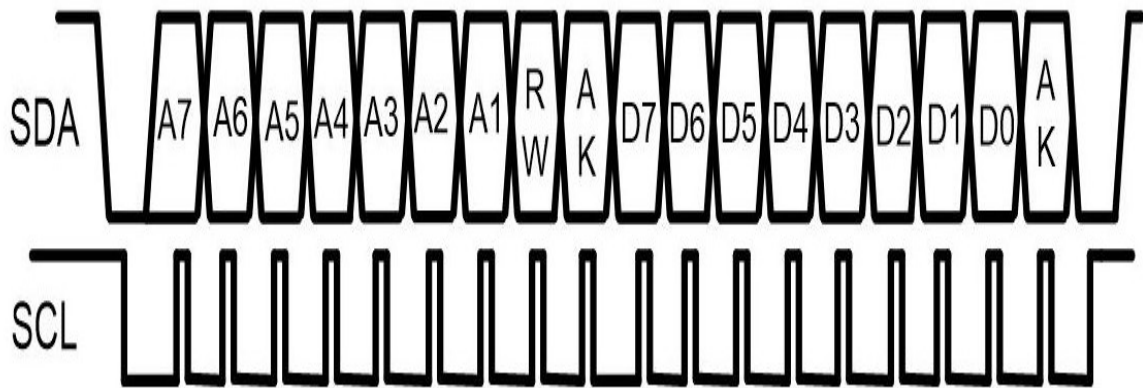


Figure 3-35. I2C Signaling Strategy

An example of writing data is shown. A Master initiates a transfer by dropping the SDA, and then lowers the SCLK. Following that, an address, Read-Write bit, and data field are sent. Acknowledgement (AK) is a response from the Slave. The clock is specified to do a full rise and fall pulse in the middle of each bit sent. With 7 bits of addressing the system is limited to 128 locations, but a more modern variant allows 10 bits of addressing.

Performance Summary: I2C	
Connection:	Multi Drop, Half-Duplex, Multi Slave, Multi Master
Data Rate:	100 Kbit/s Std, 400 Kbit/s Fast, 3.4 Mbits/s H Speed
Addressing:	7 Bits, 128 Locations & R/W
Validation:	Receive Acknowledge, No Error Check
Medium:	Wired, 2 connections
Signaling:	Ground Referenced, Active Pull Down, R Pull Up
Distance:	On PCB, Needs Common Ground
Other:	10 Bit Addressing Optional

Figure 3-36. I2C Performance Summary

The performance of I2C is limited by its data transfer rate. Also, with no error checking, the device is limited to situations where good signal integrity is a must have. Consequently, I2C is limited to local communication where all chips use the same low impedance ground, as seen on a PCB. The method fills an important need for easily implemented communication, and enjoys widespread use, with many chips utilizing I2C interfaces.

Serial Peripheral Interface (SPI)

The Serial Peripheral Interface (SPI) bus structure was defined by Motorola circa 1985. A single Master with the capability to select individual Slaves for full duplex communication is the basic structure. Many vendors have adopted the interface with some variants to suit their needs. SPI is serial data with higher speeds than a passive pullup system can provide. With a dedicated “Slave Select” pin there is no address information in the data packet. Many vendors use the SPI interface for data streaming.

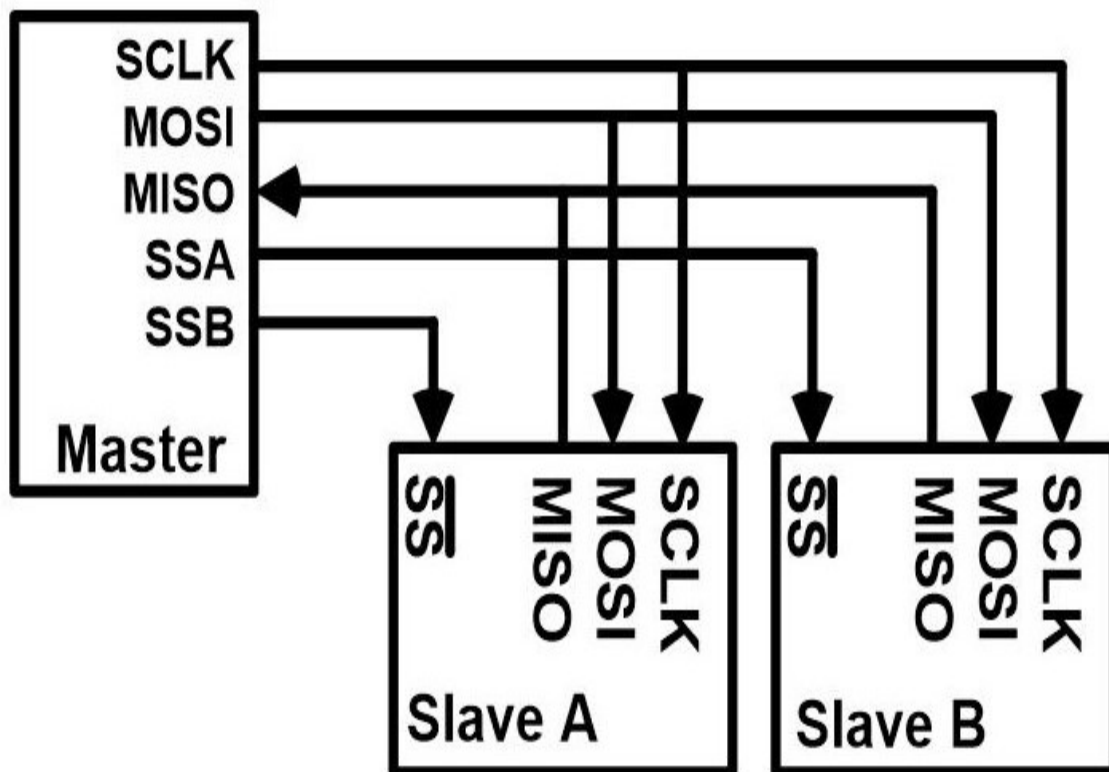


Figure 3-37. SPI Connections

SPI devices include a clock (SCLK), two data paths, MOSI (Master Out, Slave In), MISO (Master In, Slave Out). A master device requires multiple control lines to select from multiple slaves, and slave devices require an enable line, Slave Select (SS). Unless a slave has been enabled by the master, the slave maintains the shared MISO line in a high impedance state.

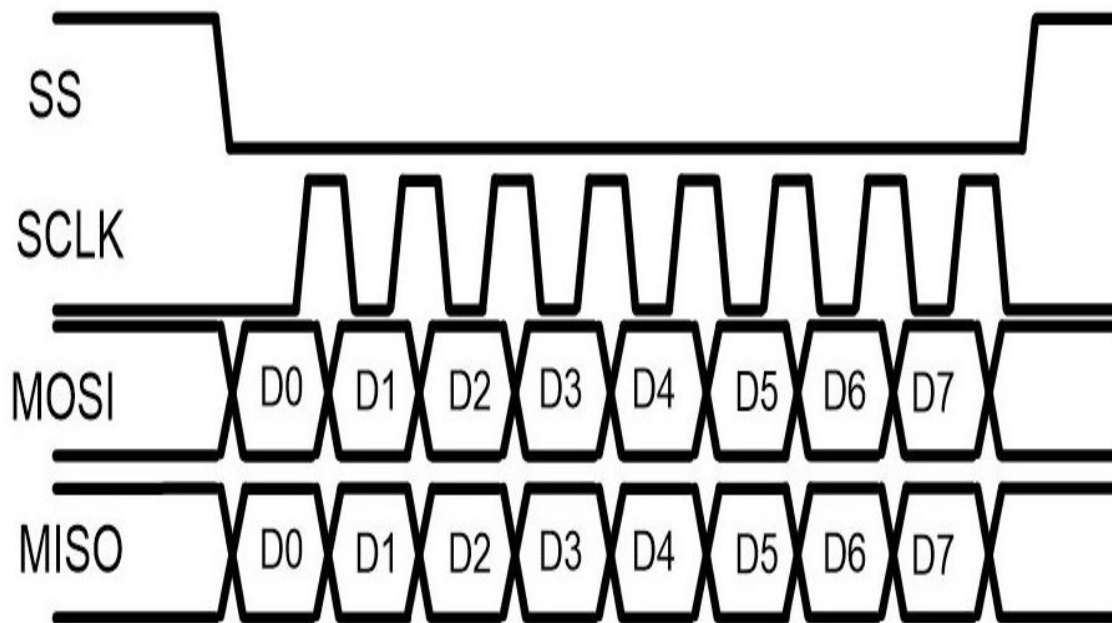


Figure 3-38. SPI Signaling Strategy

The signaling strategy is simple, the master asserts a SS connection low, and SCLK is used to transfer data in both directions. Depending on the application, the data field may be longer than shown, and some slave chips may only have a MOSI, with no MISO, functioning only as a receiver. Many variations on this standard exist.

Performance Summary: SPI (Serial Peripheral Interface)	
Connection:	Single Master, Multiple Slaves, Full Duplex
Data Rate:	10Mbits/s, is common, unique to device
Addressing:	None, Slave Select (SS) pin activates device
Validation:	No Acknowledge, No Error Check
Medium:	Wired, 4 connections
Signaling:	Ground Referenced, Active Push-Pull Drivers
Distance:	On PCB, Needs Common Ground
Other:	No formal standard, many variants exist

Figure 3-39. SPI Performance Summary

With actively switched up-down drivers SPI devices can get higher data rates than resistor pull up techniques such as I2C. However, with no error checking the method is limited to short connections on a common ground.

Single Wire Interfaces (1-wire, UNI/O)

Serial bus structures have been devised that use a single wire connection. Two of these have come into common use, namely “1-Wire” and “UNI/O”.

Dallas Semiconductor (now part of Maxim Integrated) developed the “1-Wire” interface combining data, address, and power on a single wire, referenced to the ground connection. Generally, this may be used in peripheral devices to a system such as sensors and identification devices.

Microchip developed the UNI/O bus specification which is designed for small form factor micro controller buses, single wire access to EEPROMs and other low pin count applications.

Both devices are in limited use where lowest pin count is a priority.

Serializer and Deserializer (SerDes) High Speed Data

SerDes are designed to make serial port data transfers at high rates of speed. There are multiple devices on the market with no single standard being dominant. Many vendors provide both a Serializer (Transmitter) and DeSerializer (Receiver) designed to their proprietary standard.

All SerDes have certain common features.

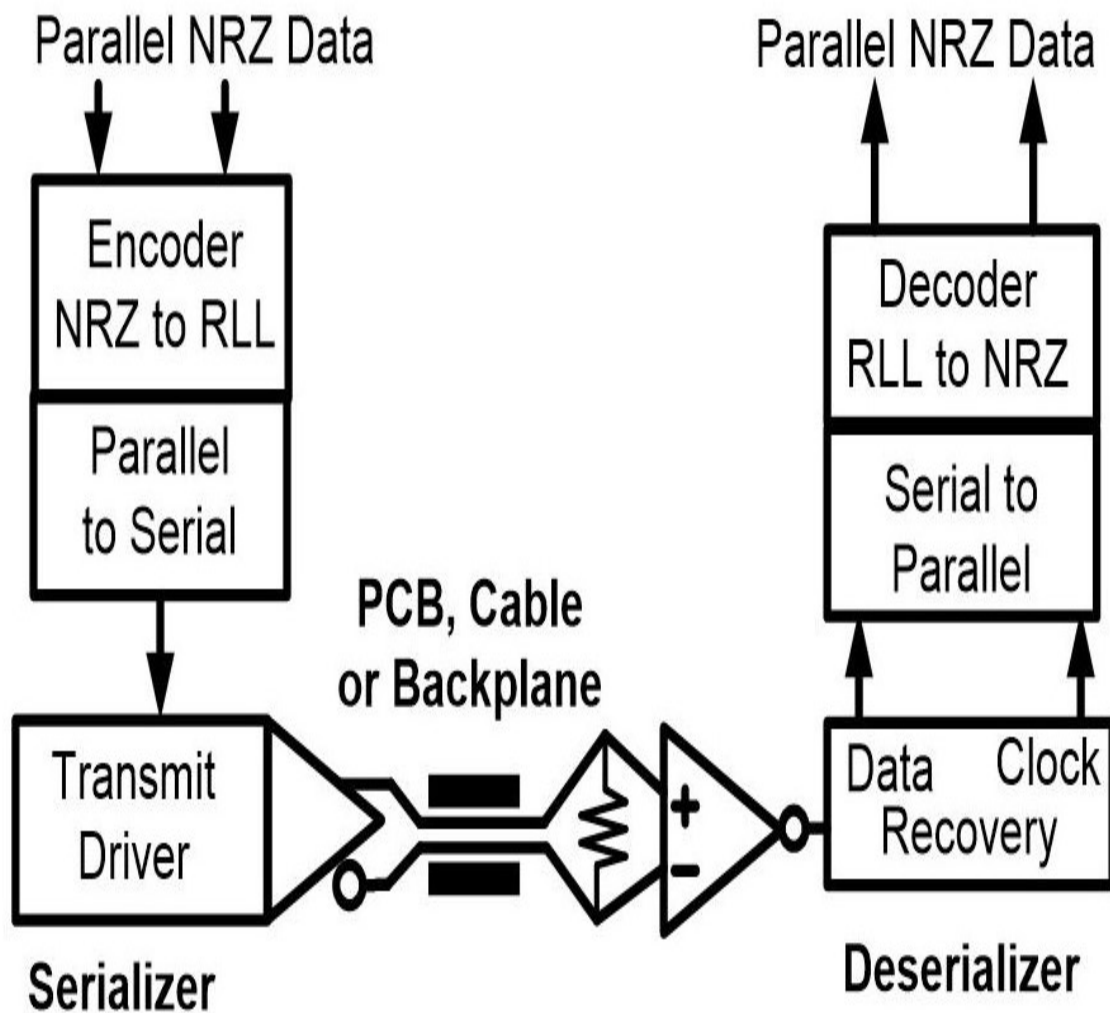


Figure 3-40. Typical SerDes Block Diagram

The Serializer and Deserializer blocks exist within separate chips, where a high speed data port is needed. SerDes use an embedded clock method as described earlier. The Serializer input is a data bus, often internal to the chip. That data bus provides a stream of NRZ data, which is then RLL encoded.

The RLL controls the transmit driver. LVDS is used as the physical strategy for transmission. Some variants use differential current steering to create a pair of signals. Some variants may use signal processing methods at transmission (pre-emphasis) and reception (equalization). The transmission

medium can be a differential stripline on a PCB, or twisted pair wires in a cable.

At the receiver, the signals are received differentially and converted into a binary asynchronous data stream. Data and clock recovery are performed, and the RLL data is decoded to NRZ synchronous data.

All SerDes use a similar architecture. At high data rates, maintaining a separate accurate clock is not readily done, so using embedded clock methods makes sense.

Performance Summary: SerDes	
Connection:	Point to Point, Simplex 2 wire
Data Rate:	Mostly < 6 Gbits/s, many variants exist
Addressing:	None
Validation:	Depends on protocol used
Medium:	Cable twisted pairs or Differential matched lines
Signaling:	LVDS or Differential Current Steering
Distance:	Depends on data rate vs. distance.
Other:	No standards here, Vendors offer chips sets to serialize/transmit and receive/deserialize.

Figure 3-41. SerDes Performance Summary

Figure 3-? is a typical architecture for one set of SerDes. Depending on what is needed, two sets can be used for duplex communication or multiple sets in parallel can be used to push data rates up. Many communications protocols use a SerDes approach to the physical layer. The specifics of a protocol to work with the SerDes will define things like a preamble, addressing, control data, data fields, error detection and related items.

Data Between Boards or Between Systems - Wired Methods

Communication between PCBs, across a cable, or between systems needs certain features to make it reliable. In the protocol, there needs to be both error checking, and handshaking methods defined to make sure the data arrived properly. The physical system needs to be differential to reject common mode noise and ground variance. It is difficult to maintain a good high frequency ground connection over distance. Widely separated systems will completely DC isolate the connection through electromagnetic or optical methods.

RS-232 functions without differential signals, but uses large ($\pm 5V$ to $\pm 15V$) voltage swings to do so. Using a 30 volt difference between logic states is a brute force technique.

RS-232 - Serial Data over Cable

RS-232 is the grandfather of serial communication, devised in 1960, when digital logic was done with discrete transistors. The original application was for teletypes with little capability to make digital decisions. Consequently, the protocol includes control signals not needed in modern methods. RS-232 was widely used on early personal computers to access printers. It is still widely in use where modern methods are lagging, such as older laboratory and industrial equipment.

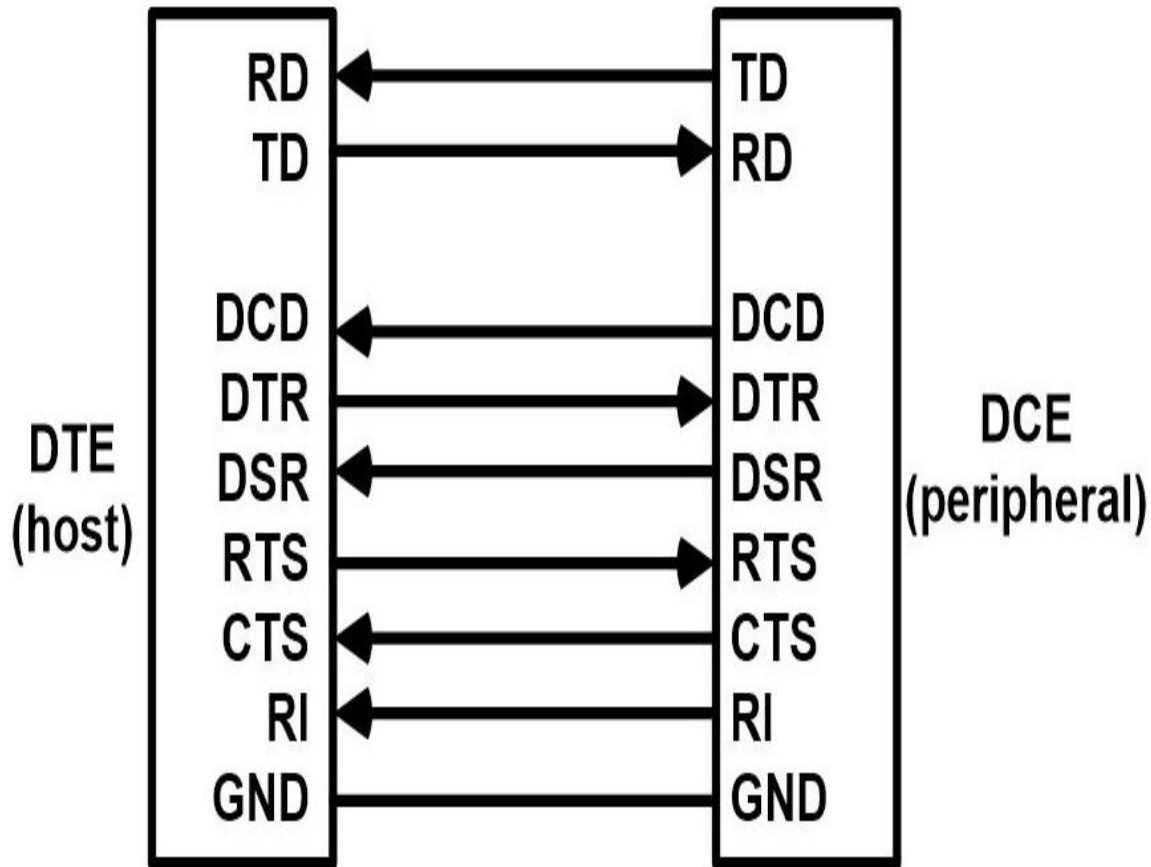


Figure 3-42. RS-232 Connections

RS-332 connects between Data Terminal Equipment (DTE, now called a “host”) and a “Data Circuit-Terminating Equipment” or “Data Communication Equipment” (DCE, now called a “peripheral”). The signals are:

- RD - Receive Data in
- TD - Transmit Data out
- RTS -Request To Send, comes from the host
- CTS - Clear To Send, comes from peripheral, ready to receive data
- DTR - Data Terminal Ready, host is ready
- DSR - Data Set Ready, from peripheral, tells host it wants to send data

- DCD – Data Carrier Detect, indicates peripheral has an active input
- RI – Ring Indicator, provides indication of incoming data

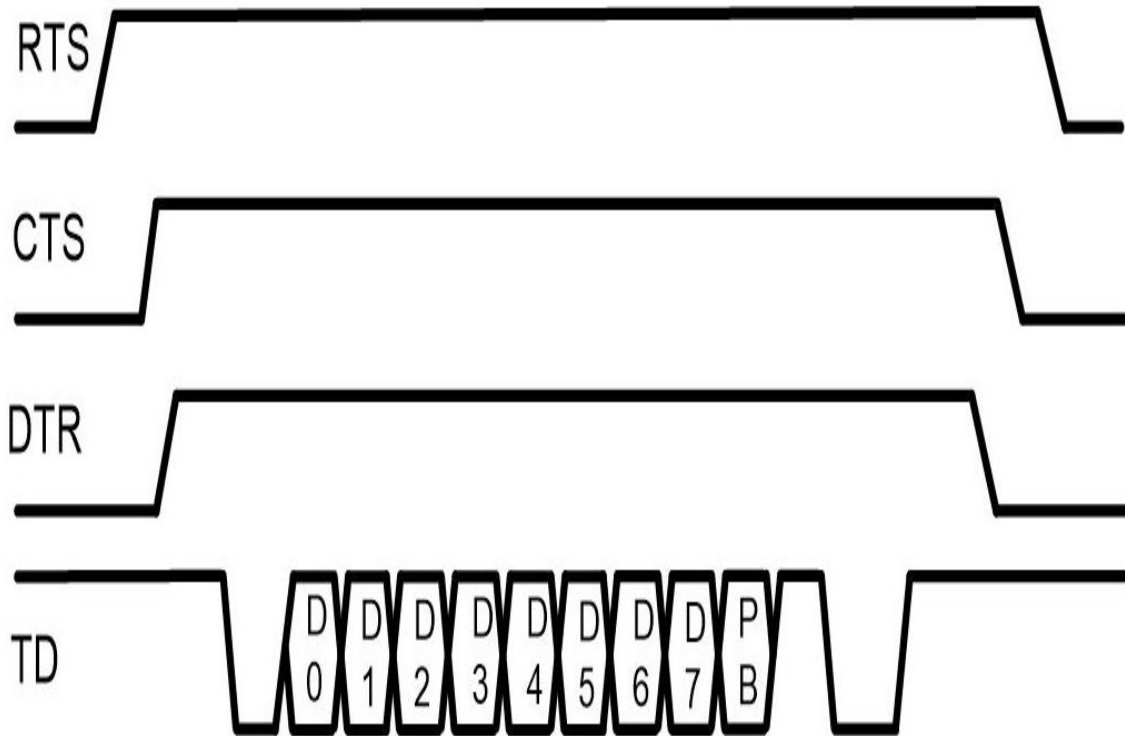


Figure 3-43. RS-232 Signaling Strategy

A typical data exchange is shown:

- Host takes RTS high telling peripheral it is ready to send
- Peripheral acknowledges with CTS high
- Host indicates it is ready with DTR high
- Data from TD goes low followed by a byte of data, and a parity bit

The first falling edge of TD is for time synchronization. The TD connection functions the same as described in the UART exchange. Similar to UART, both sides of the system must be set to the same data rate to communicate successfully, since timing is derived from the data rate, and phase location of the initial falling edge.

Performance Summary: RS-232	
Connection:	Single Point to Point, Full Duplex
Data Rate:	20 Kbits/s, maximum
Addressing:	None
Validation:	No Acknowledge, Parity Error Check
Medium:	Cable Wired, 9 pin and 25 pin variants exist
Signaling:	Ground Referenced, Wide voltage swing logic
Distance:	15 meters max, limited by cable capacitance
Other:	Obsolete, Standard exists but many “not fully compliant” variants exist

Figure 3-44. RS-232 Performance Summary

RS-232 is limited in speed and distance. Cable capacitance typically limits functionality to 15 meters. RS-232 uses ground wire referenced, large signal amplitudes in a brute force method to get a signal across.

The method is obsolete, but frequently seen in older designs. The UART variant is still in use. Neither RS-232 or UART are encouraged for new designs, since there are better methods.

RS-485 Differential Serial Data over Cable

RS-485 was defined as an industry standard in 1983. At that time, the need for differential signaling and impedance matched terminations had become evident. RS-485 is a physical layer standard where the communication protocol is not defined. Communication protocols to use the RS-485 physical layer include Modbus, Profibus, DH-485, and others. RS-485 is widely used in industrial automation and robotic control systems. RS-485 was originally specified to connect up to 32 devices, although some vendors have created interface chips that support more.

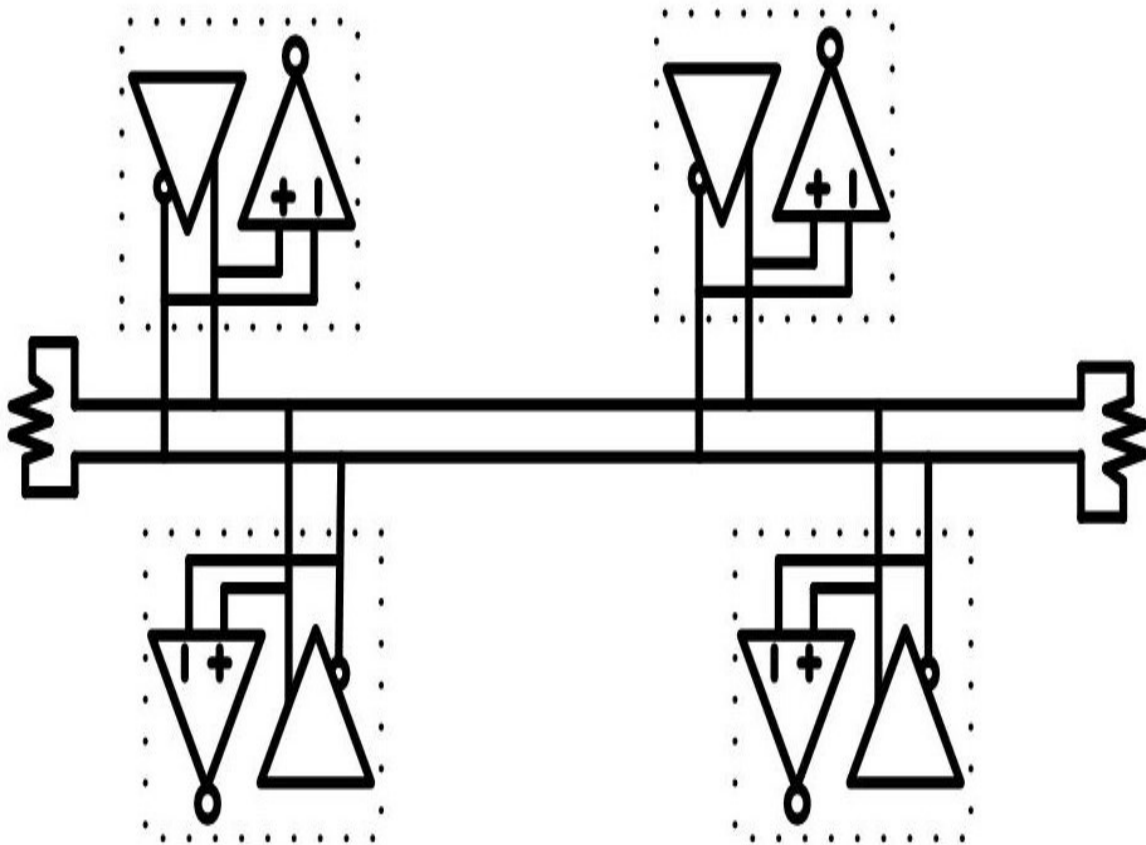


Figure 3-45. RS-485 Connections

A twisted pair interconnect makes the bus connection that is used between devices. Devices attached to the bus are daisy chained in a row avoiding transmission line stubs branching off from the bus. Termination resistance exists at both ends of the daisy chain to minimize transmission line

reflections. Essentially, a transmitter in the middle of the chain is transmitting out to two transmission lines at the same time. Signal amplitudes are specified as +/-1.5V at the transmitter and as +/-200mV (or greater) at the receiver. This allows proper functionality, even with amplitude losses in the interconnect.

Performance Summary: RS-485	
Connection:	MultiPoint, Half Duplex (2 wire) Full Duplex (4 wire)
Data Rate:	40 Mbits/s (<10 meters) 0.1 Mbits/s (1000 meters)
Addressing:	Depends on protocol used, max 32 devices
Validation:	Depends on protocol used
Medium:	Cable Wired, unshielded twisted pairs
Signaling:	LVDS, +/- 1.5V Output. Min. +/- 200mV Receiver
Distance:	1200 meters max, limited by cable resistance
Other:	Is a “physical only” standard. Protocols: Modbus, et. al.

Figure 3-46. RS-485 Performance Summary

The capabilities of RS-485 are suitable to many situations where data rates are modest, multi-point connections are preferred, and a need for

differential signal noise immunity is needed. Many vendors produce RS-485 compliant transceiver chips. Consequently RS-485 is widely used and is still being implemented into new designs.

Controller Area Network Bus (CAN Bus)

The CAN Bus was developed for communication in automobiles but is also useful elsewhere. Bosch published the first version in 1983 intending to simplify automotive wiring interconnects. Production autos using CAN Bus emerged after 1991. Virtually all cars after 2008 use CAN Bus.

In addition to automotive, CAN Bus is widely used in industrial automation, robotics, medical devices, electromechanical prosthetics, airplanes, and others. Systems using multiple uC's that need communication and coordination are well suited to a CAN Bus interface.

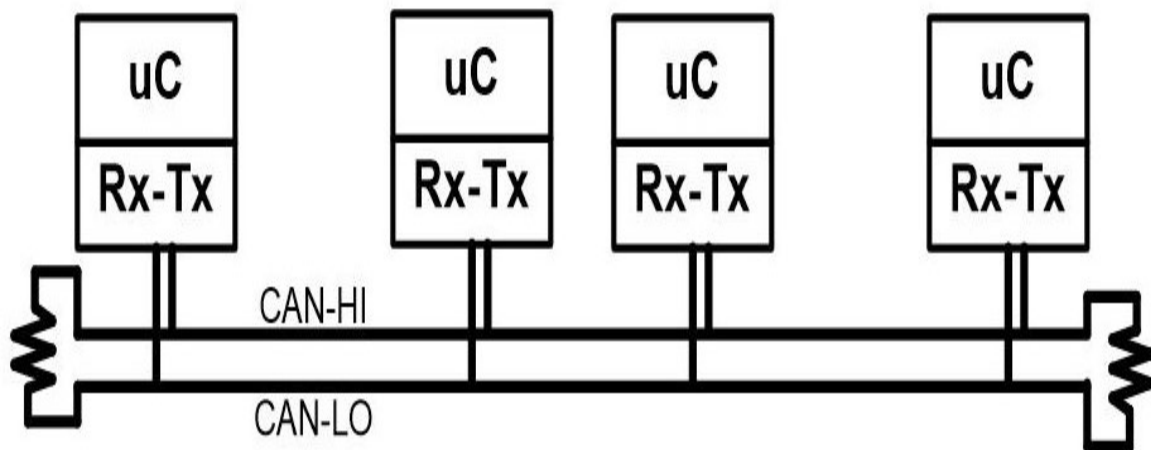


Figure 3-47. CAN Bus Connections

CAN Bus uses a multi point half duplex architecture connected by a daisy chained twisted wire pair. Signaling is differential but is not LVDS. Instead, High-Low Differential Signaling (HL-DS) is used, where a control line can be either Dominant or Recessive. The advantage of differential signaling and common mode noise rejection is still retained.

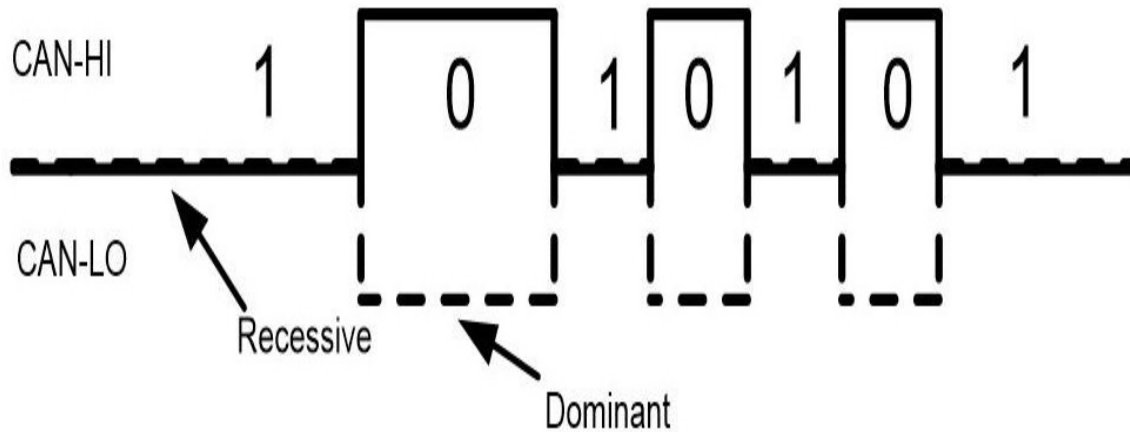


Figure 3-48. CAN Bus Signaling Strategy

In the HL-DS method, CAN-HI and CAN-LO take either a dominant or recessive state. When recessive, both signals are roughly the same mid scale voltage, which indicates logic 1. When dominant, CAN-HI is pulled up and CAN-LO is pulled down which indicates logic 0. With this, any device can force a dominant 0, over another device with a recessive 1. This is valuable in bus arbitration. When a device starts to use the bus, and puts out an address, another device can force an address change to a lower value. Address locations with lower values take precedence. Assigning addresses by priority gets faster response to select devices. In a car, the brakes take higher priority than climate control.

CAN Bus was optimized to support a distributed multi controller system with up to 30 devices on the bus. Many semiconductor vendors are providing CAN embedded drivers and software for their microcontrollers.

Performance Summary: CAN (Controller Area Network)	
Connection:	Multi Point, Half Duplex (2 wire)
Data Rate:	1 Mbits/s
Addressing:	11 or 29 Bits, max 30 devices
Validation:	Receive Acknowledge, CRC Error Check
Medium:	Cable Wired, 1 twisted pair, shielded or unshielded
Signaling:	HL-DS, 3V/1V Output. Min. +/- 200mV Receiver
Distance:	40 meters
Other:	uC friendly protocols. CANOpen, DeviceNet, CAN Kingdom

Figure 3-49. CAN Bus Performance Summary

Performance of CAN is limited to 1MBit/s and 40 meters of distance. For the target application of telemetry and coordination communication, common in cars, the data rate and distance are suitable. CAN also sees widespread use in aerospace, industrial robots, and factory automation.

Serial Data for Computer Systems

Some serial data methods were developed for use in the support of computer systems. Some remain associated with computer motherboards (Serial ATA, PCI Express) others receive use in other applications (USB, Ethernet).

Universal Serial Bus (USB)

USB was first developed in 1996 as a single cable connection for computer peripherals. USB is now the predominant method of single cord, plug and play, peripheral connection for computers, laptops, tablets and smart phones. Also, the USB cable includes power connections, leading to adoption of the USB cable for battery charging.

USB 1.0 and 2.0

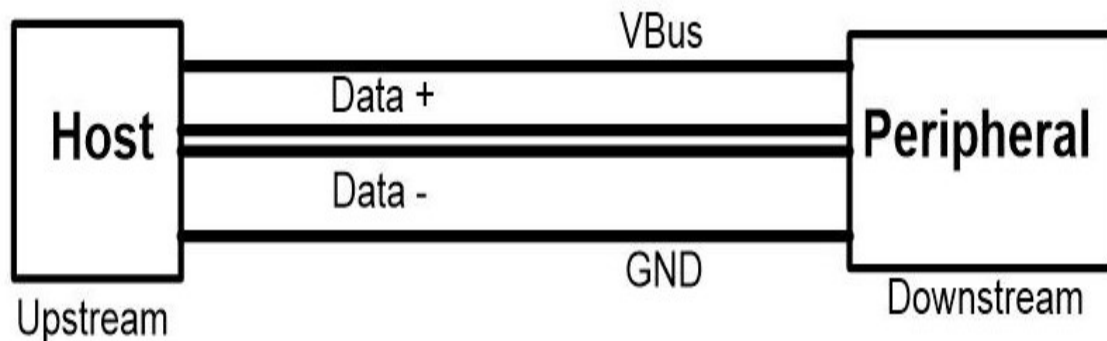


Figure 3-50. USB Connections

USB 1 and USB 2 are 4 wire (5V Power, Data+, Data-, Ground) connections. USB 3 adds two additional Data pairs for “Super Speed” data cables. USB specifications extensively define cables, power restrictions, connection design, and backwards compatibility issues to ensure reliable plug and play device compatibility.

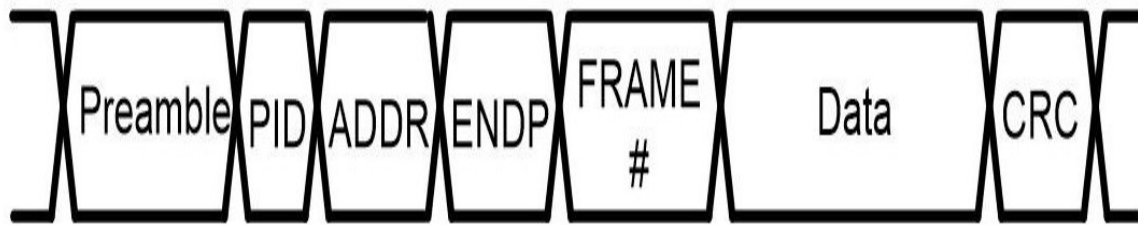


Figure 3-51. USB Signaling Strategy

Digital content of a USB transaction, called a packet, is fairly straightforward:

- Preamble – Used for receiver synchronization
- PID - Packet Identification, defines purpose of the packet; confirmation handshaking, data exchange, configuring data rate, tokens, direction of transaction, initialization, etc.
- ADDR - Target address, for a maximum of 127 devices.
- ENDP – Endpoint, provides supplemental information to support the PID. Depending on PID value, the ENDP can have unique functions.
- FRAME# - Frame Number used as a sequencing tool to properly order multiple packets.
- DATA - Data Field can be from 1 to 1024 bytes in length.
- CRC - Cyclic Redundancy Check (CRC) Error checking for the data field.

Data bytes are transmitted LSB first, MSB last.

Performance Summary: USB 1.0, 2.0, 3.0	
Connection:	Point to Point, Half Duplex (1.0, 2.0)
Data Rate:	1.5M, 12M, 480M, 5G bits/s versions
Addressing:	7 Bits, 127 devices max
Validation:	Handshaking protocol, CRC Error Check
Medium:	4 wire cable, single twisted pair, Power, GND
Signaling:	LVDS with added series resistance
Distance:	5 meters (1.0, 2.0), ~3 meters (3.0)
Other:	USB 1,2 are 4 wire, USB 3,4 are 9 wire, USB 3.1-4.0 are up to 40Gbit/sec. Can hot swap.

Figure 3-52. USB Performance Summary

USB devices must be user friendly. Cable lengths are limited, but easy plug and play scenarios, automatic configuration, and hot plugging are well supported. The price paid for an easy end-user experience is the complexity of the support software. USB software can be a significant part of the design effort.

Serial Advanced Technology Attachment (SATA)

SATA is a specialty computer interface for mass storage devices. Magnetic hard disk drives (HDD), Solid State Disks (SSD), and optical media (DVD, CD) use this interface when internally installed. Stand alone mass storage devices have adapted to a USB interface for the easier plug and play setup.

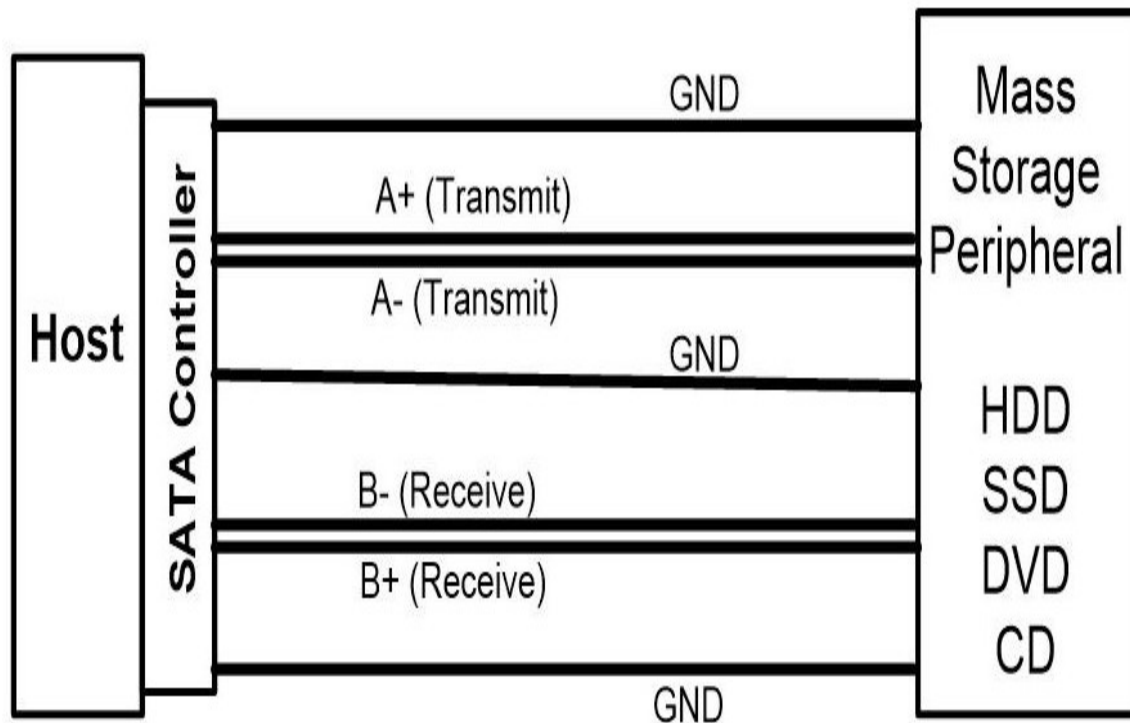


Figure 3-53. SATA Connections

SATA has 5 wires, two (A,B) differential twisted pair connections and three grounds. Power is independently supplied. Although separate transmit/receive paths are available the interface is not full duplex. Half duplex capability is specified, due to the inability of a HDD to write and read simultaneously. Address format reflects HDD technology using the internal format of a HDD. Address format includes “Head” (recording head selection) “Cylinder” (track cylinder) and “Sector” (which rotational “wedge” area).

Performance Summary: SATA (Serial AT bus Attachment)	
Connection:	Point to Point, Half Duplex
Data Rate:	150M, 300M, 600M bits/s versions
Addressing:	Magnetic Media Address: Head, Cylinder, Sector
Validation:	CRC Error checking
Medium:	7 wire cable, two twisted pairs, Power, GND
Signaling:	LVDS
Distance:	1 meter
Other:	Specialty interface for mass storage devices.

Figure 3-54. SATA Performance Summary

SATA is a specialty interface that designers will need only when dealing with a mass storage device.

Peripheral Component Interconnect Express (PCIe)

PCIe originated in 2003 as a computer bus expansion standard. As a serial bus interface, it replaces the older parallel bus approaches, PCI and PCI-X. PCIe can be thought of as multiple SerDes channels. In PCIe, a full duplex

communication “lane” uses two SerDes channels, one for transmission and one for reception.

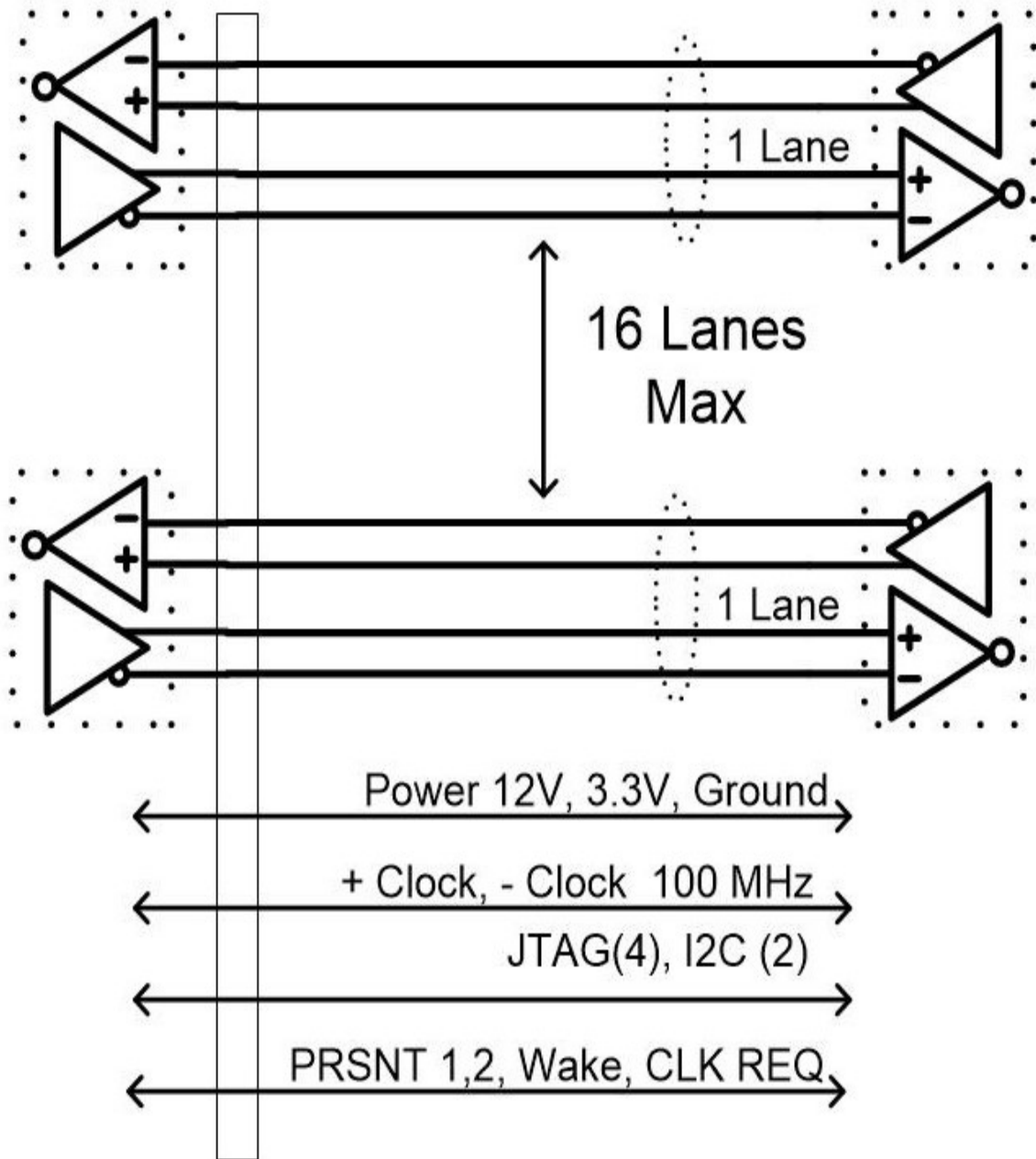


Figure 3-55. PCIe Interconnect Structure

Up to 16 lanes can be included where high data rates are needed. A 100 MHz LVDS clock is provided across the bus connection. This clock serves as the reference for the receiver’s clock and data recovery. Other features

include 12V and 3.3V power, multiple ground connections, JTAG interface for testing, and SMBus (aka I2C) for configuring the remote device. Supplemental control connections to quickly determine how many lanes are present (X1, X4, X8, X16 variants) are provided by PRSNT#1, PRSNT#2 pins. WAKE control and CLKREQ complete the needed controls. PCIe performance is quite fast, but distance limited.

Performance Summary: PCI Express (PCIe)	
Connection:	Point to Point, Full Duplex
Data Rate:	Up to 4G bits/s per lane, 16 lanes max, ~64 GB/s
Addressing:	32 or 64 Bits
Validation:	Acknowledge protocol, CRC Error Check
Medium:	18 to 82 pin connector, data across PCB
Signaling:	LVDS for data and clock
Distance:	50 cm
Other:	Fully specified standard for signals, connectors, protocol, testing.

Figure 3-56. PCIe Performance Summary

Version 1(2003) is 250 MB/s per lane, version 5 (2020), is 3938 MB/s per lane. Using 16 lanes of version 5, provides 63 GB/s throughput. High speed comes with a distance restriction of 50 cm. PCIe is designed to interface plug in cards to a motherboard backplane. If high speed data on a cable is needed, conversion to a cable friendly format is the solution. Both Ethernet and USB are suitable, depending on needed data rate and distance.

The PCIe specification fully defines electrical, mechanical, and protocol criteria to ensure plug in compatibility between vendors. PCIe capability is included on microprocessors, FPGAs, digital video interfaces, and others, where high speed data is needed

Ethernet

Ethernet is the LAN workhorse of computer systems. Early versions used coaxial cable and some modern variants use optical fiber. The bulk of Ethernet connections are done using differential signaling over twisted wire pairs. Here the focus is on the ubiquitous twisted pair Ethernet embodied in the 10Base-T, 100Base-TX, 1000Base-T variants.

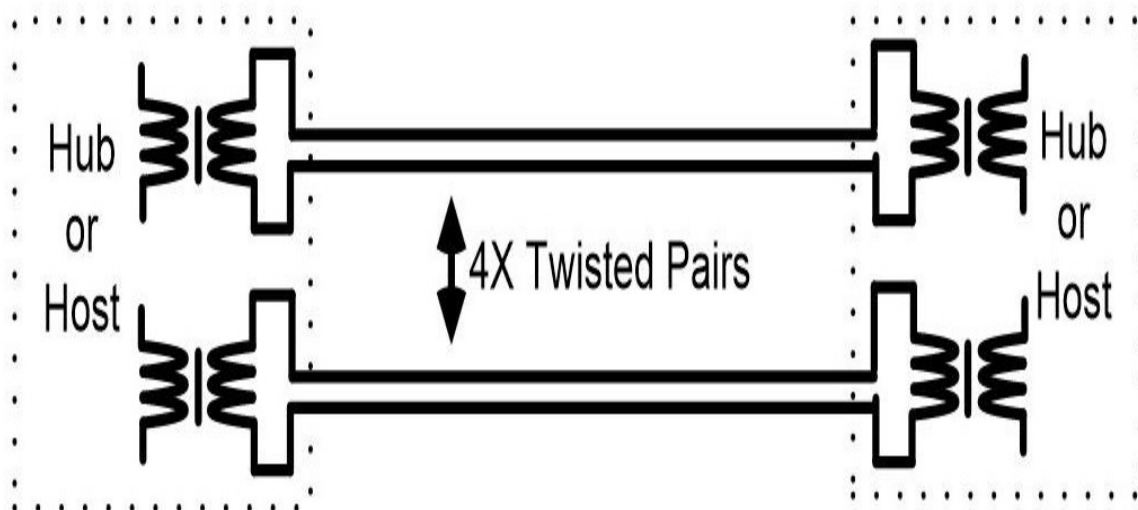


Figure 3-57. Ethernet Connections

Ethernet over twisted pairs is done with 4 differential signals, using 8 wires. The signaling method changes with the data rate. An important part of the Ethernet connection is the magnetic coupling at each end. Using a

transformer and common mode choke at both ends eliminates common mode signals. This gives up to 1500 volts of DC isolation and immunity to voltage differences between the grounds at each end. Ethernet uses point to point connections, avoiding signal corruption issues associated with multi drop systems. Connections between cables are avoided using active circuits that perform signal regeneration and retransmission thus isolating Ethernet cables from each other.

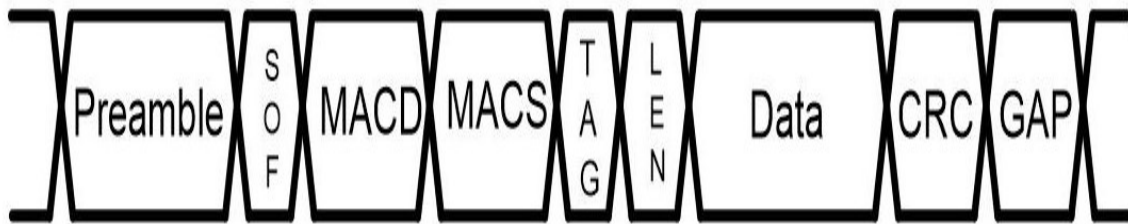


Figure 3-58. Ethernet Signaling Strategy

Information contained in an Ethernet message is referred to as an Ethernet Frame. Information in the Frame must support a system where source and destination are not always the two devices on each end of the cable.

Contained in the Frame are:

- Preamble – Used for receiver synchronization.
- SOF - Start of Frame, indicates end of preamble and start of the address field.
- MACD - Media Access Control Address Destination, 48 bit address of the final destination of the Frame.
- MACS - Media Access Control Address, Source, 48 bit address of the original source of the Frame.
- TAG - Information about priority and protocol to be used for the Frame.
- LEN –Length of the data field.
- DATA - Data field, 46 to 1500 bytes long.

- CRC - Cyclical Redundancy Check, error checking code. Also known as a Frame Check Sequence.
- GAP - Required idle time before the next Frame can be sent.

Performance Summary: Ethernet over Twisted Pair Copper	
Connection:	Point to Point, Full, Half Duplex (4, 8 wire, versions)
Data Rate:	10, 100, 1000 Mbits/s versions
Addressing:	48 Bits, Media Access Control (MAC) addressing
Validation:	Receive Acknowledge, CRC Error Check
Medium:	Cable Wired, 2 or 4 twisted pairs
Signaling:	DS, Manchester(10), MLT-3 (100), PAM-5 (1000)
Distance:	100 meters
Other:	Other Ethernet variants exist. Herein covering 10Base-T, 100Base-TX, 1000Base-T

Figure 3-59. Ethernet Performance Summary

Ethernet capability covered here includes up to 1GB/s variants. There are Ethernet variants over copper that will function at 5GB/s, 10GB/s, and 40GB/s.

Wireless Serial Interfaces

Wireless methods of making a data connection are commonplace and easy to implement. Many options exist to fit specific needs.

Wi-Fi

Wi-Fi came into use in 1997, and quickly became the foremost method to obtain internet access without a cable. Designing Wi-Fi into a system is now a modular approach with complete Wi-Fi modules requiring only a digital interface. Designers need very minimal knowledge of RF design to successfully include Wi-Fi in a system.

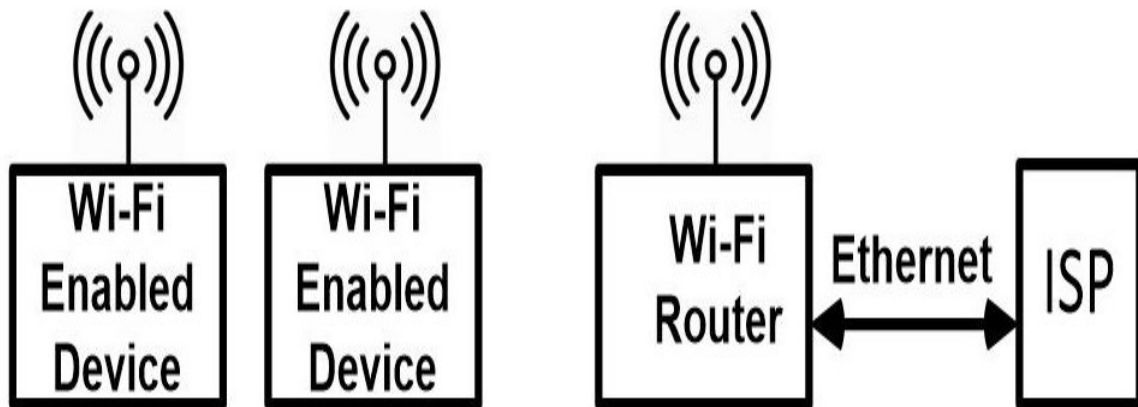


Figure 3-60. Wi-Fi Typical Application Scenario

Conceptually, the application of Wi-Fi is straightforward. An internet service provider (ISP) is connected to a Wi-Fi Router. Devices with Wi-Fi capability (Laptops, tablets, smart phones) can obtain internet access through the wireless link.

Performance Summary: Wi-Fi	
Connection:	Multi Point Wireless
Data Rate:	1- 54 Mbits/s versions are common
Addressing:	48 Bits, Media Access Control (MAC) addressing
Validation:	Error Check and Acknowledge
Medium:	900 MHz, 2.4 GHz, 5 GHz, 5.9 GHz are common
Signaling:	100mW typ. QPSK (rev B), 16/64 QAM (rev A, G)
Distance:	50 meters, omnidirectional antenna,
Other:	Multiple updates with different performance of distance and data rates. Above are typical.

Figure 3-61. Wi-Fi Performance Summary

Wi-Fi operates on frequencies suitable to line of sight communication (ISM bands at 2.4GHz and 5GHz in the USA). At 100 mW of power, distance is limited to roughly 50 meters. Directional antennas and special care with obstruction free transmission paths have yielded longer distances, but for general purpose communications, with obstructions, interference issues, and omnidirectional antennas, 50 meters is a suitable guideline.

Bluetooth (BT)

BT was developed for battery powered, close proximity, wireless connections. BT has been widely adopted. Some examples:

- Computers: wireless keyboards, mice, game controllers, and speakers.
- Cell phones: wireless headsets, ear buds, smart watches, speakers, linking phone to car.
- Others: medical body sensors, machines needing unrestricted motion, hearing aids, handheld remotes free of alignment sensitivity

All of these devices predominantly use BT or BT LE methods.

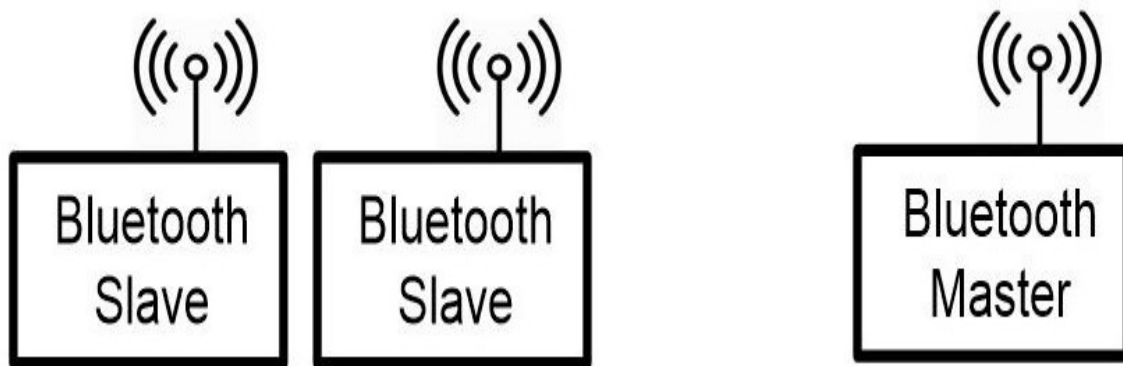


Figure 3-62. Bluetooth Typical Application Scenario

BT works in a master slave configuration with up to seven slaves. BT is suited to creation of a PAN (Personal Area Network)

Performance Summary: Bluetooth	
Connection:	Wireless, 1 master, 7 slaves max.
Data Rate:	1Mbits/s (1.0) 3Mbit/s (2.0)
Addressing:	48 Bits, Media Access Control (MAC) addressing
Validation:	Error Check and Acknowledge
Medium:	2.4 GHz
Signaling:	Frequency Hop. GFSK (1.0) DQPSK/8-DPSK (2.0)
Distance:	10cm (1mW), 10m (2.5mW), 100m (100mW)
Other:	Optimized for battery powered wireless, with rates up to 3Mbit/s.

Figure 3-63. Bluetooth Performance Summary

BT operates in the unlicensed ISM band at 2.4GHz. Using variable power output from 0.5 mW to 100 mW BT distance is roughly limited to 100 meters (high power) and 0.5 meters (low power). For things like cell phone headsets data rates are sufficient to support high quality audio, and small form factor battery life with 1-2 days of general use.

Many devices don't need a data rate to support audio, or have a low amount of on time associated with transmitting data. This motivated the creation of

Bluetooth LE

Bluetooth Low Energy (BTLE)

Bluetooth LE (BTLE, also known as Bluetooth 4.0) has similar applications as BT in closely spaced networks. BTLE is optimized for items that require low average data rates and intermittent data transfers. By maintaining a low power sleep mode, with brief wake times for communication, BTLE can achieve a battery life of years, while BT devices are measured in days. BTLE in versions 5.1 (introduced 2017) and beyond includes the capability for mesh networks, (aka Bluetooth Mesh, or BT-Mesh) catering to the Internet of Things (IoT) market.

Performance Summary: Bluetooth Low Energy V4.2	
Connection:	Wireless, 1 master, # slaves application dependent
Data Rate:	125 kbits/s, 1MBit/s, 2MBit/s
Addressing:	48 Bits, Media Access Control (MAC) addressing
Validation:	Error Check and Acknowledge
Medium:	2.4 GHz
Signaling:	Frequency hop GFSK
Distance:	10cm – 50m (1mW – 10 mW)
Other:	Optimized for low power for extended battery life. Not suitable for duplex audio.

Figure 3-64. Bluetooth LE Performance Summary

Similar to Wi-Fi, using BT or BTLE in a device is a black box approach requiring only a digital interface to get on the air. Designers need minimal RF knowledge to include BT/BTLE into a system.

ZigBee

ZigBee (introduced 2005) is designed for low data rates, low power sleep mode, brief communication bursts, mesh network capability and over 2

years of battery life. ZigBee is optimized to work in a mesh network of many devices.

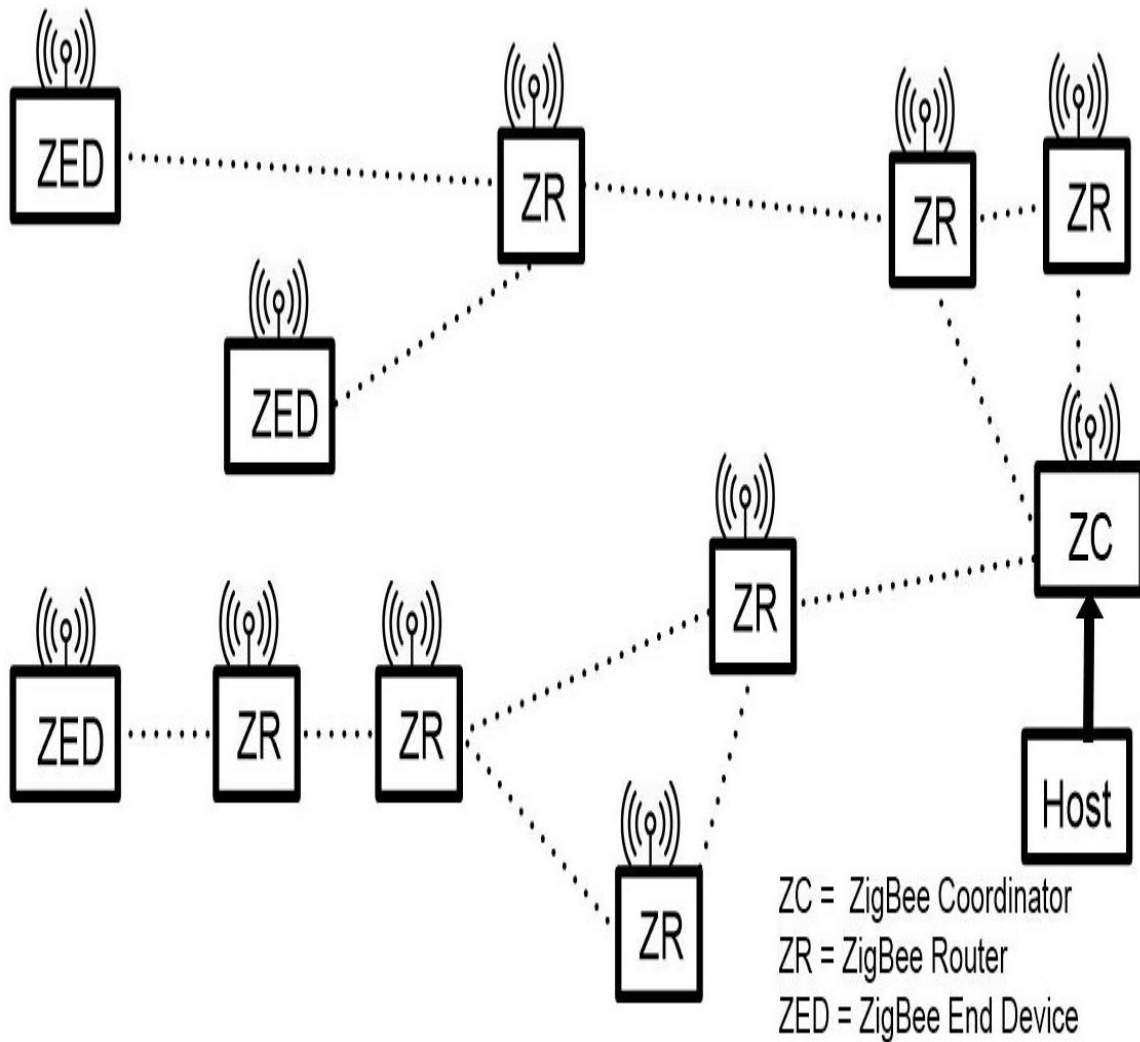


Figure 3-65. ZigBee Typical Application Scenario

Devices connect wirelessly through multiple paths of communication. Home automation networks of distributed devices such as alarm system sensors, electronic locks, smart appliances, and other IoT devices are possible.

Every ZigBee mesh network has one coordinator (ZC), and can have multiple routers (ZR) and multiple end devices (ZED). ZR devices can pass signals through and serve to extend the network, ZED devices can connect

to the mesh but not extend the network. The ZC orchestrates the system. For a device to be ZigBee certified a battery life of 2 years needs to be met.

Performance Summary: ZigBee Mesh Network	
Connection:	Wireless, 1 coordinator(ZC), multi ZR, multi ZED
Data Rate:	250 kbits/s
Addressing:	64 Bits, Media Access Control (MAC) addressing
Validation:	Error Check and Acknowledge
Medium:	2.4 GHz
Signaling:	OQPSK
Distance:	10m-100m (1-100 mW), fixed location devices
Other:	Optimized for low power requires 2 year battery life to be certified.

Figure 3-66. ZigBee Performance Summary

ZigBee devices use a 64 bit MAC address similar to a MAC 48 bit Ethernet address, allowing a unique address. The largest number of the devices allowed on a mesh network is 64,000.

Z-Wave

Z-Wave is a wireless mesh network similar to ZigBee. Instead of the 2.4 GHz ISM band, it operates at 865-930 MHz, depending on country. (908 MHz in USA) Other differences include a limitation of 4 hops in the mesh connection and a maximum of 232 devices in the mesh.

Adaptive Network Topology (ANT)

ANT started as a wireless link for sports and fitness sensors. BTLE and ANT devices compete for market share where ultra low power consumption and low data rates are a priority.

Infrared Data Association (IrDA)

Infrared light is commonly used as a digital data link for wireless remote controls and communication where a device must be electrically isolated. Infrared is limited to distances under 2 meters and is directionally sensitive. Handheld remote controls for home entertainment devices are still predominantly infrared due to low cost. Some remote controls are migrating to non infrared methods (BTLE is common) to avoid directional sensitivity problems.

Fiber-Optic Data – Go Fast, Go Far

Data over copper has limits of both speed and distance. As of 2020, PCI Express maximum data rate is 63 GB/s but is limited to 50 cm. Ethernet over twisted pairs goes to 10GBit/s but is limited to 100 meters. There will be faster capability after this publishes, but to go fast and go far requires fiber-optic links.

An in depth look at fiber is beyond the scope of this book, but a quick performance summary is useful.

Performance Summary: Data over Optical Fiber	
Connection:	Point to Point, Multiple channels per fiber
Data Rate:	50G – 500G bits/s per channel
Addressing:	Depends on protocol used
Validation:	Depends on protocol used
Medium:	Optical fiber, multiple channels in each fiber
Signaling:	QPSK, QAM, OFDM and others are used
Distance:	>100 Km
Other:	Multi channels/fiber & multi fibers/cable are done, Cables with >100T bits/s are done. .

Figure 3-67. Data Over Optical Fiber Performance Summary

Optical fiber will go far (kilometers) and go fast (Terabytes/s). Consequently, fiber is the backbone of the internet. If you need high performance data links, optical fiber methods need to be investigated.

JTAG – PCB Access for Test and Configuration

JTAG (Joint Test Action Group) is a digital serial access port for PCB testing. Many digital chips include a 4 pin JTAG access port for testing after PCB assembly. JTAG ports on multiple chips connect together, allowing access to all chips with one external port connection. Controller chips, FPGA's and EEPROM's are often programmed through the JTAG port after board assembly.

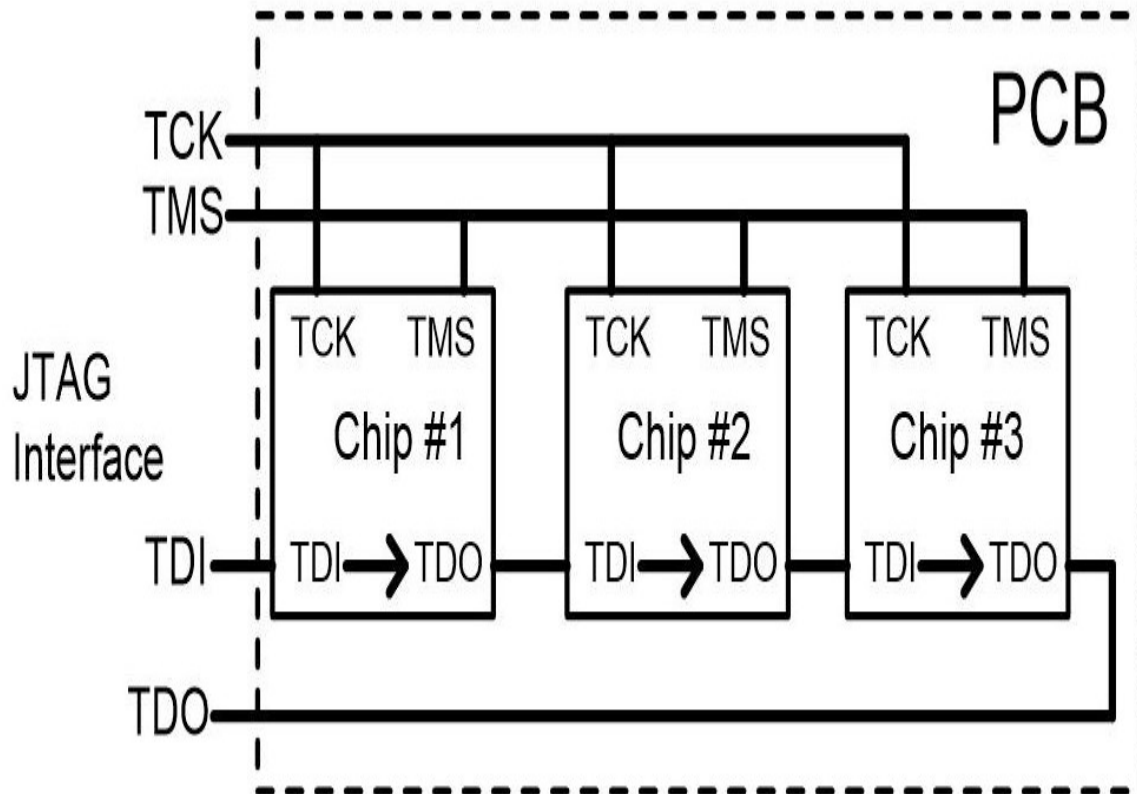


Figure 3-68. JTAG Typical Application Scenario

JTAG Signals:

- TDI -Test Data In, Serial Port
- TDO -Test Data Out, Serial Port
- TCK - Test Clock allows clocking of the serial data path
- TMS - Test Mode Select

TMS and TCK control an internal state machine (TAP Controller) that controls the JTAG interface. The TAP Controller allows loading internal registers or pipelining data out through the TDO and into the TDI of the next device. In addition to internal control, JTAG is used to control boundary scan testing. This controls both the digital output ports, and senses the digital input ports of an IC. This allows testing of interconnects between chips on the PCB.

Performance Summary: JTAG	
Connection:	Wired Daisy Chain (2,4,5 connection variants)
Data Rate:	No standard, typically under 10 MHz clock
Addressing:	Serial pipeline
Validation:	None
Medium:	4 connection implemented in PCB traces
Signaling:	Ground referenced logic
Distance:	On PCB
Other:	For PCB testing, boundary scan and code downloads to digital IC's (IEEE Standard 1149.1)

Figure 3-69. JTAG Performance Summary

JTAG is commonly implemented using the 4 wire variant. The speed of the JTAG chain is limited by the slowest chip in the chain. Generally, TCK is under 10 MHz. JTAG is valuable for automated production testing and programming of on board devices. All PCBs should have a JTAG port to make boards testable and programmable. Including a “Design for Test” strategy is an important part of any design.

Essential Concepts – Conclusions and summary

The important points of this chapter are:

- Digital signals may have signal integrity problems
- Slow-Soft-Short-Networks are generally reliable
- LVDS provides many advantages for fast or distance applications
- Network placement and routing affects performance
- Comparing rise time signal length to connection length is a guide to lumped vs. distributed networks and terminated transmission line need
- Significance of time skews are guided by what percentage of the clock period they encompass
- Physical length path balancing for time skews is often unnecessary, designers need to be quantitative to determine need
- Serial port communication is the dominant connection method
- For chip to chip communication on a common PCB use I2C(SMBus) and SPI
- SerDes methods are useful for fast data
- Distance communication needs common mode rejection, error detection and hand shaking validation
- For cable communication between systems use: RS-485, CAN Bus, Ethernet

- USB is the predominant host to peripheral interconnect
- Computer interfaces internally with PCIe and SATA
- Wireless interconnect options: Wi-Fi, Bluetooth, ZigBee (and others) offer easy design-in methods
- Fast and Far Data requires optical fiber methods
- Include a JTAG interface in any PCB for test and programming access

After reading this chapter, informed decisions can be made on how to get data between two points. For brevity, many specific details had to be left out. Knowing what's needed can start the implementation process. The needed details are available in many of the documents listed in the "Further Reading" section.

Further Reading

Digital Design Techniques

"High Speed Digital Design, A Handbook of Black Magic", 1993, H Johnson, M Graham, Pearson, ISBN 978-81-317-1412-6

"High-Speed Digital System Design – A Handbook of Interconnect Theory and Design Practices," 2000, S. Hall, G. Hall, J. McCall, John Wiley & Sons, ISBN 0-471-36090-2

"Digital Signal Transmission", 1992, Cambridge University Press, C.C. Bissell, D.A. Chapman, ISBN 0-521-41537-3

"Dealing with High-Speed Logic", Analog Devices, MT-097 Tutorial Rev 0, January 2009

"Signal and Power Integrity Simplified" 3rd edition, E. Bogatin, 2018, Pearson Education, ISBN-13, 978-0-13-451341-6

"Comparing Bus Solutions", F. Aliche, M. Feulner, F. Dehmelt, A. Verma, G. Becke, Texas Instruments, Application Report, SLLA067C, Rev. August 2017

“Digital Transmission Lines – Computer Modeling and Analysis” K. Granzow, 1998, Oxford University Press, ISBN 0-19-511292-X

LVDS

“LVDS Owner’s Manual, Including High-Speed CML and Signal Conditioning”, Fourth Edition, Texas Instruments, 2008

“LVDS Application and Data Handbook”, Texas Instruments, SLLD009, November 2002

I2C and SMBus

“I2C Manual”, Philips Semiconductor, Application Note, AN10216-01, March 2003

“The I2C Bus Specification,” Version 2.1, January 2000, Philips Semiconductors

“I2C Bus Specification and User Manual” NXP Semiconductors, UM10204 Rev. 6, April 2014

CAN Bus

“CAN Bus - ISO 11898-1:2015” <https://www.iso.org/standard/63648.html>

“Introduction to the Controller Area Network (CAN),” Texas Instruments, Application Report, SLOA101B, Revised May 2016

“Controller Area Network Physical Layer Requirements,” Steve Corrigan, Texas Instruments, Application Report, SLLA270, January 2008

“Controller Area Network (CAN) Implementation Guide,” Conal Watterson, Analog Devices, Application Note, AN-1123 Rev. A

RS-485

“The RS-485 Design Guide” T. Kugelstadt, Texas Instruments, Application Report, SLLA272C, Rev. October 2016

“RS-485 & Modbus Protocol Guide”, Tyco Electronics, Rev 6 July 2002

Ethernet

“A Beginner’s Guide to Ethernet 802.3”, R. Newhaus, Analog Devices, Engineer to Engineer Note, EE-269, Rev 1 June 2005

SPI

“SPI Block Guide”, Motorola Inc., Version 03.06, rev. February 2003

JTAG

JTAG Testability Primer, IEEE Std 1149.1, Texas Instruments, 1997

SATA

“Serial ATA – High Speed Serialized AT Attachment”, SerialATA Workgroup, Revision 1.0a, January 2003

USB

“Universal Serial Bus 2.0 Specification”, Revision 2.0, April 2000

“Universal Serial Bus 3.0 Specification”, Revision 3.0, November 2008

Special Interest Groups for Protocols, Products, and Interfaces

Wi-Fi Special Interest Group can be found at: <https://www.wi-fi.org/>

Bluetooth Special Interest Group can be found at:

<https://www.bluetooth.com/>

ZigBee Alliance can be found at: <https://zigbeealliance.org/>

Z-Wave Special Interest Group can be found at: <https://www.z-wave.com/>

ANT Special Interest Group can be found at: <https://www.thisisant.com/>

Infra Red Special Interest Group can be found at: <https://www.irda.org/>

Chapter 4. Power Systems, Method, and Distribution

A NOTE FOR EARLY RELEASE READERS

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

This will be the 4th chapter of the final book.

If you have comments about how we might improve the content and/or examples in this book, or if you notice missing material within this chapter, please reach out to the editor at arufino@oreilly.com.

Providing system power can be a simple matter, but for complex systems it can turn into a significant design issue. Power and ground are common to an entire system and can be a smoky failure point, a common path for noise into the entire system, or a path where the power demand in one place affects the performance someplace else. Safe AC power, safe failure scenarios, regulatory compliance issues, suitable noise levels on power, controlled up/down power cycles, and stable power under highly variable loads are all part of a reliable system. Understanding options available among power converters and how to configure the power system are the goal. Designing an AC to DC converter from scratch is rarely necessary, with many off the shelf converters available that are easily implemented into a system. After initial AC/DC power conversion, additional items such as: local regulation, multiple voltages for different circuits, high frequency bypassing, selective device decoupling, power monitors, and power cycling controls are commonly needed.

Several source power scenarios exist: disposable battery, battery with charging capability, external AC/DC converter, or AC/DC converter inside the system. All of these are looked at, and how they are similar or different.

Remember That: Split Phase AC Mains Power

What does AC mains power consist of (Figure 4-1)? In the USA, at the residential level, the AC power grid feeds a local stepdown and distribution transformer that creates a “split phase 120/240 system” output. What reaches a typical wall outlet (120 VAC 60Hz in the USA, and 50Hz elsewhere) is a three wire system, of Line (aka Hot), Neutral, and Earth Ground. Neutral is the center tap from the secondary in the AC mains stepdown transformer.

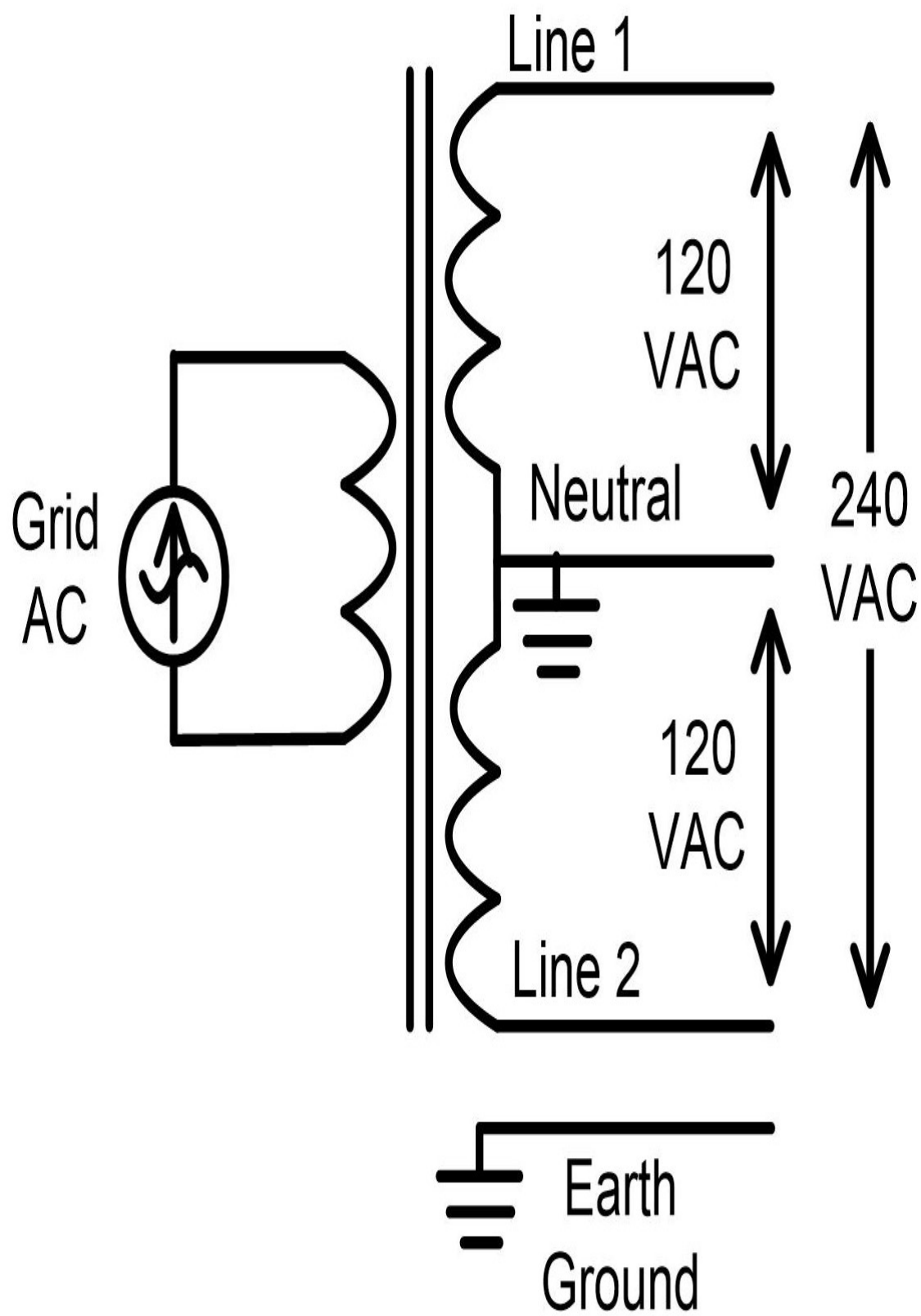


Figure 4-1. AC Split Phase Power, Stepdown from Grid

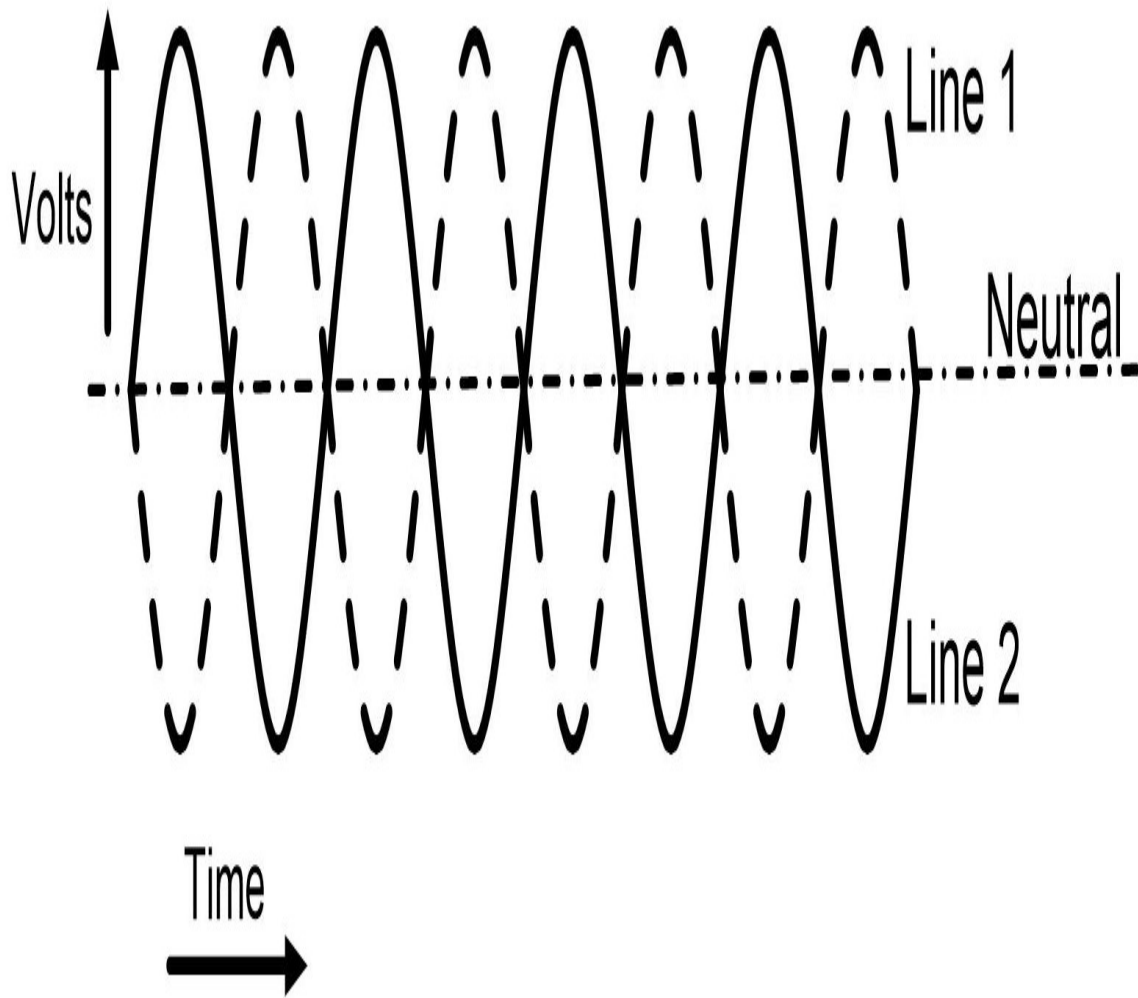


Figure 4-2. AC Split Phase Power, Signals

Normally, a current passes through Line and Neutral with zero current in the Ground. Under normal operation, Neutral and Ground should be at roughly the same voltage. This 120 VAC is not a symmetric differential signal (Figure 4-3) only the Line signal has a sinusoid. However, when both ends of the secondary of the AC mains transformer are used (240VAC in the USA) the output voltage is a symmetric differential signal centered at the Neutral signal. Devices with an AC mains input exclusively use the Ground connection for safety. The ground serves as a conductive enclosure

connection that guarantees the outside of device remains at ground potential. Neutral and Ground are only connected together back at the AC Mains distribution transformer. Prior to 1960, residential wiring was 2 wires only, without the ground safety, which produced unsafe devices in many failure scenarios.

AC power safety, Defining the Problem

When bringing AC power directly into a device, the implementation needs to be carefully considered in the design. Getting things wrong with AC power can lead to deadly devices, and smoke filled outcomes. Most devices now use low voltage DC power internally, provided from an external AC/DC converter and allows most designers to remain blissfully unaware of high voltage product safety requirements. Many older (prior to 1960) electrical products were inherently unsafe or could become deadly with a single fault failure. Metal enclosures, (prior to the widespread use of plastics) and a lack of a ground safety connection, made for a deadly combination. Thankfully, consumer electronic products have a fairly short life, so most of those devices are no longer in use.

Low Voltage DC Safety

Designing with power inlets under 50V are classified as “Extra Low Voltage” (ELV) and that simplifies things a lot, putting safety issues largely in the hands of the designer of that AC/DC converter.

ELV devices are subdivided into “Separated ELV” (no ground return path used) and “Protected ELV” (a ground earth safety is present). The safety standard, IEC 62368-1 (Safety Requirements for Audio/Video, Information & Communication Technology Equipment) has become the standard for ELV safety. For ELV devices, safety requirements are very minimal, with some form of overcurrent protection and a non-conductive enclosure generally sufficient.

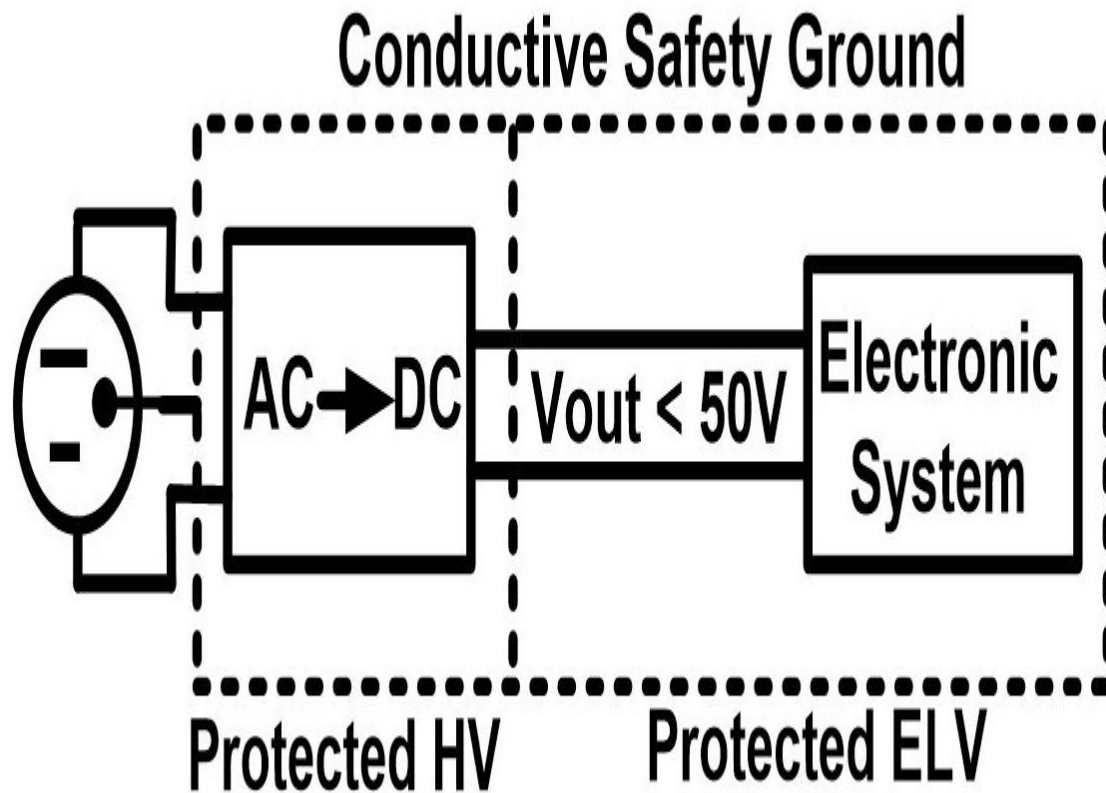


Figure 4-3. Protected HV with Protected ELV DC

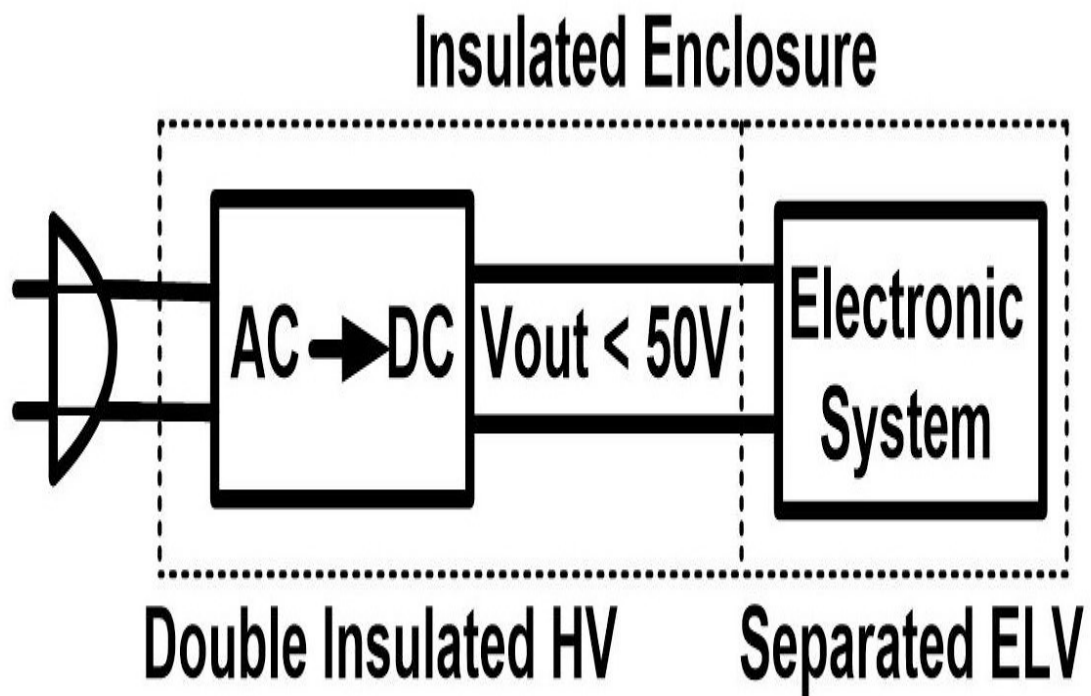


Figure 4-4. Double Insulated HV with Separated ELV DC

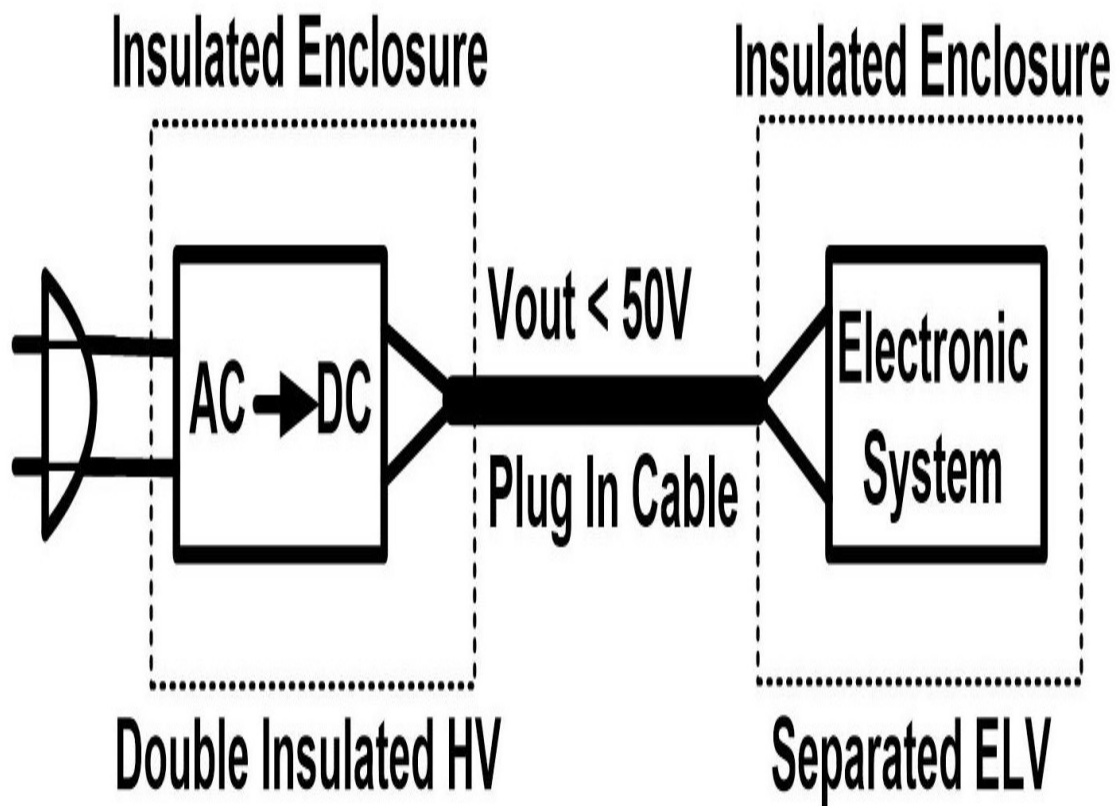


Figure 4-5. Strand Alone Double Insulated HV with Separated ELV DC

Safe Failure Methods and Single Fault Safe Scenarios

Normal operational safety is the start, but when something breaks, power systems also require a safe failure method. The “Single Fault Safe” concept is simple: Any element of the system can fail in an open circuit state, closed circuit state, or contact any other element in the system and not cause danger to the user.

Introduction of a ground safety return path into modern power wiring allowed for a safe system in most single fault situations, where the grounded enclosure, in conjunction with a fused “Line” connection, could keep a device safe. According to the international standard IEC-61140, devices that use a metal chassis and enclosure structure which is attached to

a ground safety connection are defined as Class I devices. Their failure safety is dependent upon the presence of a properly connected external ground.

Some devices have only Line and Neutral connections, and meet safety requirements using a “Double Insulated” structure. For these devices, the internal wiring and circuits are within a secondary insulating enclosure. In this manner, the device can have a single fault failure without exposing users to high voltage contact. Double insulated devices are considered Class II devices and the safety redundancy of the device is not dependent on an external ground. Consequently double insulated devices are popular in electric power tools, due to remaining safe when connected to electrical power with questionable ground paths, as frequently happens in construction environments.

ELV systems are Class III devices under the IEC 61140 regulatory standard. If the AC/DC is placed inside the product, the whole product needs to meet Class I or Class II safety requirements of a grounded enclosure, or double insulated configuration. Using an AC/DC converter externally circumvents that restriction and is a popular approach to the problem.

Overcurrent Protection Methods and the Weakest Link

Overcurrent protection is the most basic form of system protection. What are the options here? There are four common ways to turn things off in an overcurrent situation: fuses, circuit breakers, PTC devices, and active circuit methods.

Fuses are commonly used for their simplicity and low cost. However fuses can be too slow, they require manual replacement to restart the system, and they are not terribly accurate. In some situations you would like to control both current and time needed for the fuse to do its job. However, how much overcurrent and the activation time are interactive, so don't expect fast and accurate from a fuse. Nonetheless, the fuse is often fast and accurate

enough. That, combined with, low cost, simplicity, and reliability, and it is still the go-to solution in many cases. If the fuse is expected to blow only when something major on the PCB fails, a soldered in surface mount fuse can be used with minimal cost and excellent reliability. Fuses don't need to be replaceable in many cases and socket mounting is unnecessary.

A circuit breaker suffers many of the limitations of a fuse, with speed and accuracy being "good enough" for most situations. In order to gain the convenience of a re-settable device, the penalty is paid in cost and size. A similar rating circuit breaker to a fuse can be considerably more expensive to implement. Also, circuit breaker reliability can suffer in hostile environments due to internal corrosion issues if they are not hermetically sealed. Generally, circuit breakers are called for when the designer expects frequent overcurrent situations, due to the application.

The polymeric positive temperature coefficient device (widely known as a PTC or a PPTC) can be considered a self resetting fuse. They warm up, due to overload current, and the device goes into a high resistance mode, limiting the current until the device cools off, and then returns to a low resistance state. The devices do have some less than ideal internal on/off resistance but the PTC serves well for situations where auto reset is desirable. The PTC has limitations of accuracy and response time, as all thermally triggered devices do. Generally, the PTC can be included on the PCB for a low cost design solution.

Active circuit protection for overcurrent detection and shutdown comes in many variants. The basic concept is a set of detection circuits for current and/or voltage to determine if the applied power is outside specified limits, which triggers suitable circuit controls to shut down or limit performance. These systems can be fast and accurate if needed. The price paid here is complexity and cost. Frequently, this approach is unneeded and one of the simpler methods is suitable. Many switching AC/DC converters do have active circuit methods in their design, in order to protect power circuits that can self destruct faster than a fuse can respond. Seek out this feature when selecting AC/DC converters. Look for information that indicates the device

is “self protecting” or “overcurrent protected” or “current limited” in the specification.

Whatever overcurrent protection used, the system needs to make sure that the “weakest link” is the actual protection device. If the PCB traces melt at 1A of current and the protection device opens at 2A, something was not designed properly. Also, the parameters for the trip protection point needs to consider both min-max sustained and startup surge currents. As a simple guideline, setting a fuse value at twice the maximum current seen in normal circuit operation should be suitable. Also, systems may have high current and low current sections where a failure in the low current section requires a lower value protection shutdown for that part of the system. A common design mistake is using a single overcurrent shutdown in a situation where a low power part of the system self destructs since it is powered in parallel with the high current part of the system.

Thermal monitoring of high power devices can be useful, with protection control built into the system. For some devices this provides another layer of safety and capability to stress devices to their thermal limits without self destruction. Older devices used thermocouples, but silicon based thermal monitors with digital interfaces to a local controller are a more modern approach. Many modern high performance microprocessors include this within the IC.

AC to DC conversion, 60 Hz Transformers

Very few electronic devices use AC power directly. The majority of devices use DC power, and most modern electronics use 5V and lower power sources. Systems with DC motors or other high power devices may need higher voltages. Due to this, the first converter in your system will be an AC/DC converter, with sufficient voltage to support the high voltage devices.

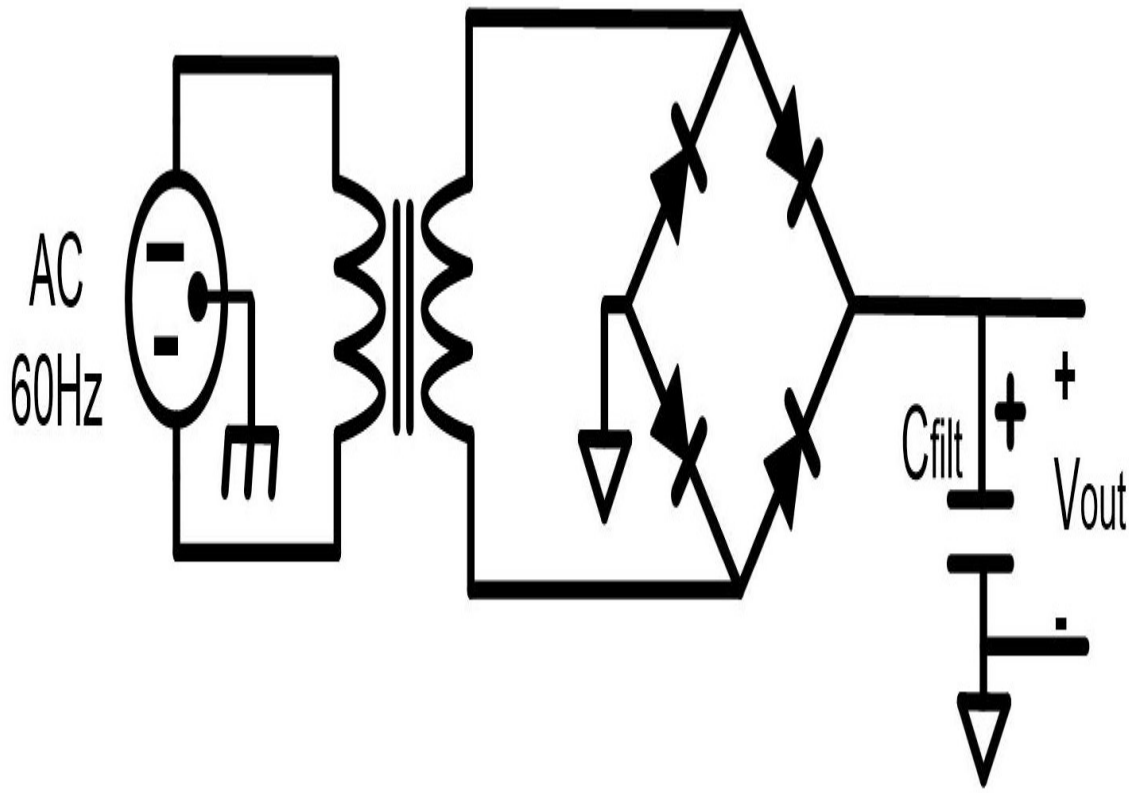


Figure 4-6. Classic 60 Hz AC to DC With Iron Core Transformer

The classic approach to AC/DC conversion uses a transformer at 60 Hz. (or 50 Hz, depending on country) This seems simple; step down the AC voltage, put the reduced AC voltage through a diode bridge, and filter the rectified AC for a DC output. The idea is simple but certain pieces tend to be both big and expensive. The problem here is the low frequency of AC mains power, because a 60 Hz transformer will be both bulky and heavy. For the same power rating, the size of a transformer reduces as the frequency goes up. In addition, the output filter capacitor has to accurately maintain the output voltage under high current loads between the rectified peaks of the 60 Hz waveform. This necessitates the use of a large electrolytic capacitor filter. If you can increase the frequency, you can reduce the size of both transformer and filter capacitor. The aviation industry uses 400 Hz AC power in order to take advantage of this, but now there are better methods.

AC to DC conversion, Off Line Switchers

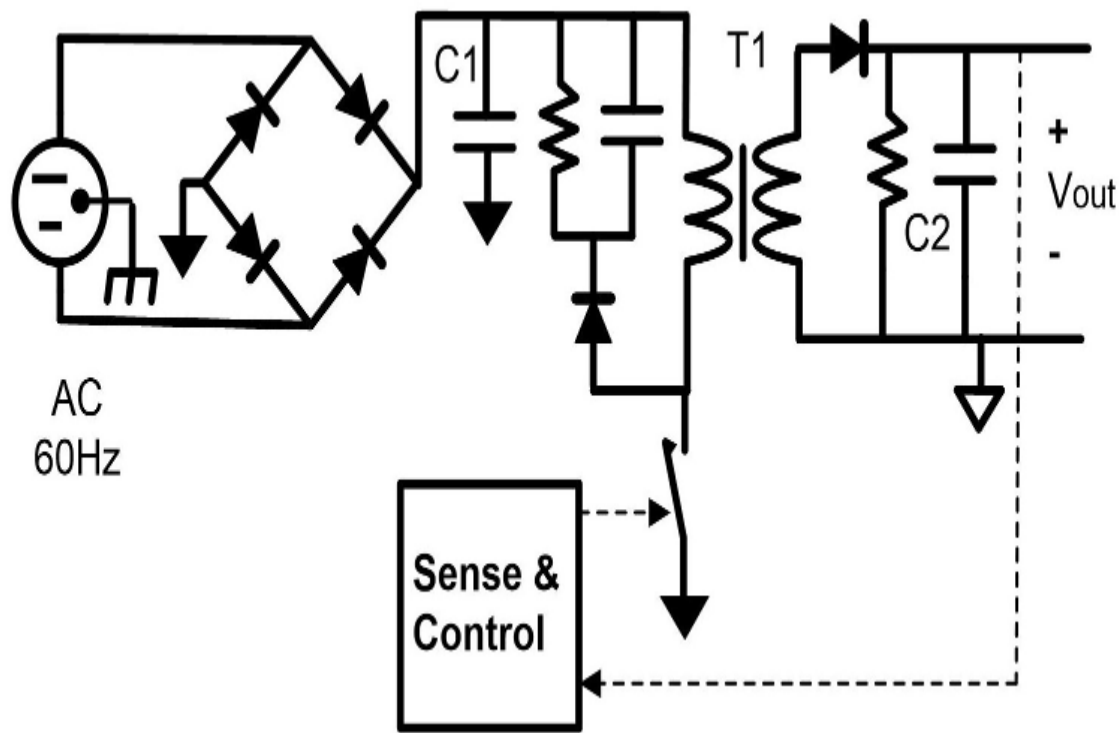


Figure 4-7. AC/DC With Off Line Switcher

The limitations of the 60 Hz AC to DC converter are circumvented with a switched mode conversion method commonly known as an “off line switcher” (Figure 4-?). This approach rectifies the AC line voltage, and uses C1 to maintain a high DC voltage where ripple and accuracy are not critical. That DC voltage is rapidly pulsed across the primary of T1 with a switch. The switching frequency is much higher than 60 Hz, with 10 KHz to 1 MHz being commonly used. Although appearing more complex, several advantages are apparent. Using a higher frequency, allows for a much smaller transformer (T1) and output filter capacitor (C2). The high voltage DC at C1 can vary, so C1 does not need to be very large. This method is used in virtually all AC/DC converters today, with small form factor implementations illustrated by the numerous USB and cell phone chargers that sit readily on an AC power plug footprint. A broad selection of AC/DC converters, using offline switching methods, are available. The

60 Hz method is much larger and more costly to implement. Due to this, the 60 Hz stepdown transformer is largely obsolete.

The AC/DC converter generally provides a single DC voltage and is commonly configured as the greatest DC voltage that the system needs. Other power regulators, for DC/DC conversion, are often needed to complete the power needs of the system.

Multi PCB Systems Need Local Power Regulation

The impedance of power and ground wires, combined with dynamic current loads will create both ground bounce and power sag. A design can reduce this effect with shorter (less inductance) and thicker (lower resistance) wires, but not eliminate it. Consequently, a multi-PCB design must deal with both ground and power voltage variance between PCBs. Using separate regulators on each PCB to create the locally used voltages needed is the widely used solution (Figure 4-?). Systems with multiple circuit boards, can use a common AC/DC converter to distribute a “raw” DC voltage. This DC voltage should have “headroom” namely, several volts larger than the minimum required by each of the satellite PCBs.

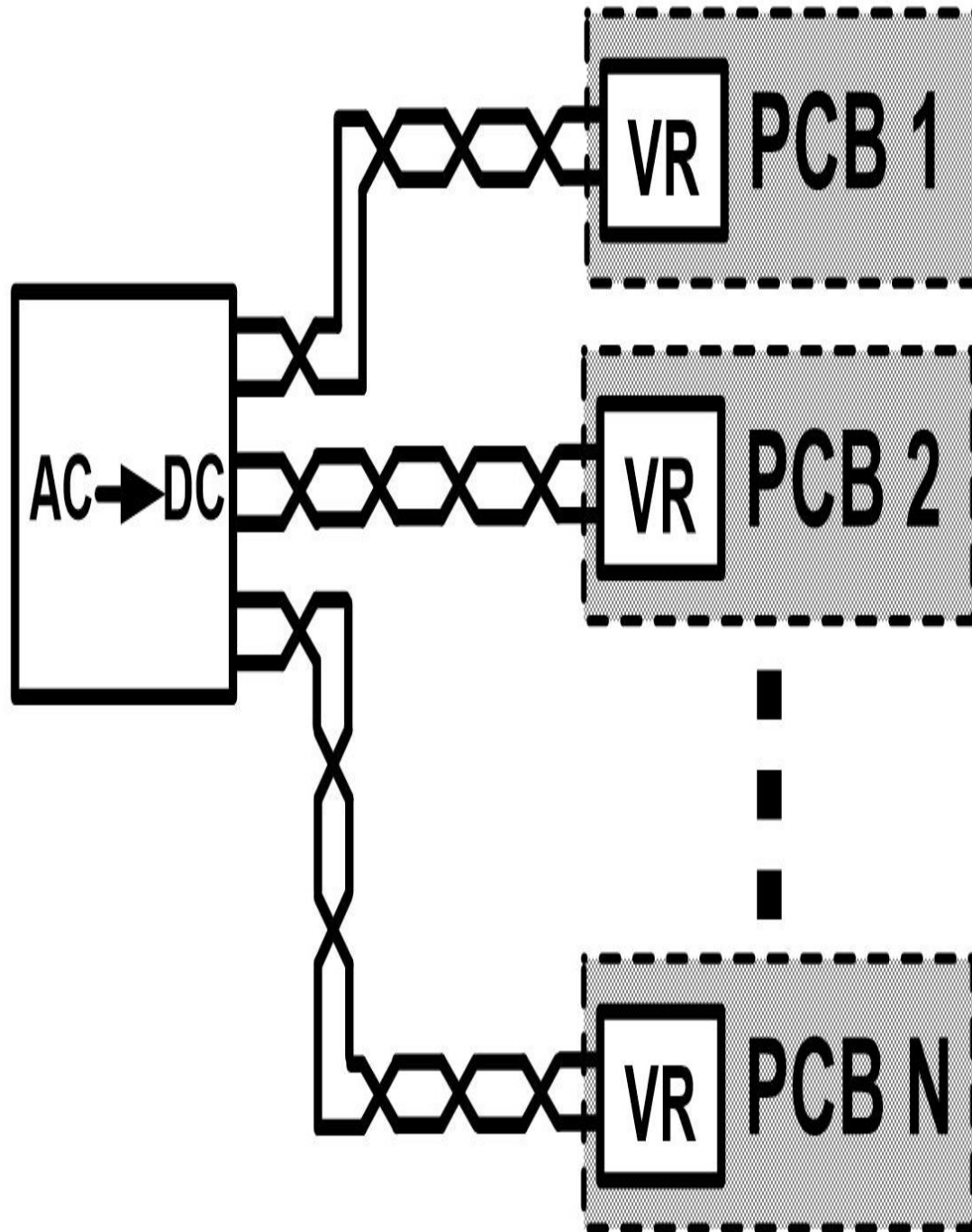


Figure 4-8. Multi PCB Systems with Local Voltage Regulation

With the added voltage headroom, the effects of ground bounce and voltage sag can then be removed by the local voltage regulator, and each PCB now

has the ability to tolerate power and ground variance at the power input. Many industrial systems distribute raw DC power at 36V, 24V or 12V depending on what the satellite PCB's require. Using a DC voltage, created off the PCB, without a local regulator, will never be consistent or noise free, and frequently causes failures during EMC regulatory testing. Regulate your power supply locally on the same PCB where you use it.

Power regulators, DC/DC Conversion, Linear vs. Switching

Power regulators can be divided into two groups, linear regulators and switching regulators, which use distinctly different methods to provide a consistent voltage output.

Linear regulators include Emitter Follower Regulators, and Low Drop Out (LDO) regulators. Switching regulators include Buck converters (voltage reducing) Boost converters (voltage increasing) and Buck-Boost converters, which can reduce or increase voltage. Other Switching Mode Power Supplies (SMPS) configurations exist as well, but these are the ones most commonly used.

Linear Regulators, Conceptual

$$P_{\text{burned}} = (V_{\text{in}} - V_{\text{out}}) / I_{\text{out}}$$

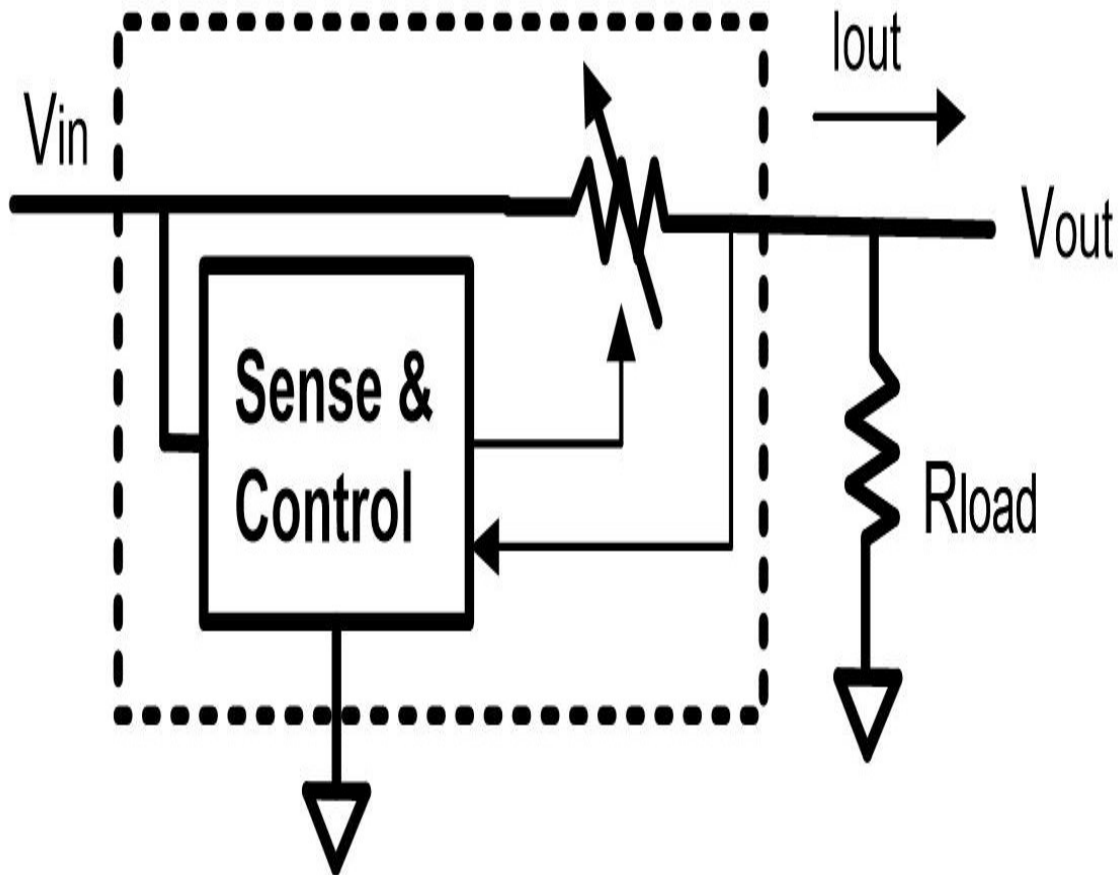


Figure 4-9. Conceptual Linear Regulator

Conceptually, a linear regulator (Figure 4-?) uses adaptive voltage division. As R_{load} changes, a sense/control circuit adjusts the regulator to maintain V_{out} . Here, a power transistor is serving as a variable resistor. The important thing to recognize is that this method creates an output voltage by burning energy inside the regulator. Due to that, when using a linear regulator, you cannot expect efficient use of energy. Also, power dissipated (P_{burned}) by the device can be quite high depending on the current (I_{out}) and the voltage drop ($V_{\text{in}} - V_{\text{out}}$) across the regulator. The power dissipated needs to be carefully considered so that the regulator does not thermally fail.

The primary advantage of linear regulators is that they are electrically quiet, in both radiated emissions, and create a fairly quiet output voltage. The big disadvantage is that they are energy inefficient and often not viable for large input to output voltage drops.

Emitter Follower Regulators vs. LDO

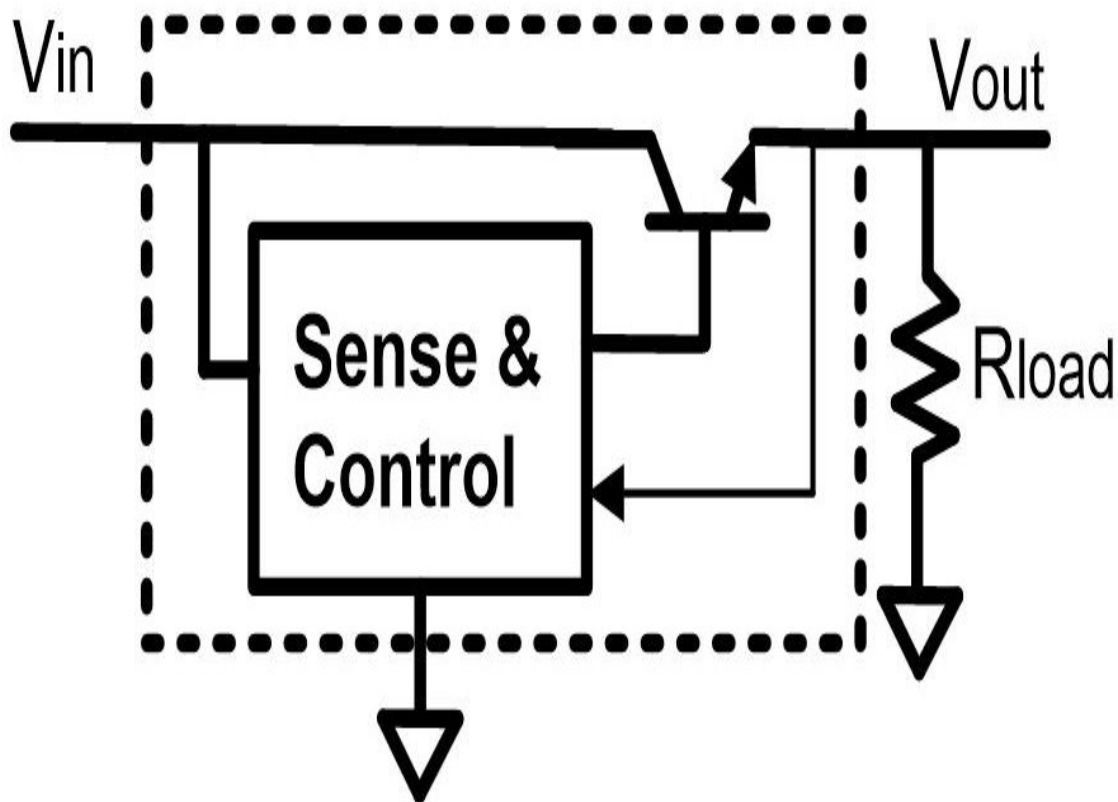


Figure 4-10. Linear Voltage Regulator, Emitter Follower

The oldest version of the linear regulator is the emitter follower configuration (Figure 4-?). A control and sense circuit applies an appropriate control voltage to the base-emitter junction of a transistor, to produce the output voltage. The transistor serves as the top half of the voltage divider, and is commonly implemented with a NPN bipolar device. The control voltage goes up, and the output voltage follows along. Due to the internal circuit design, the input voltage needs to be typically 2 volts

greater than the output voltage to properly function. The emitter follower regulator configuration is widely used in the 7800 series (7805, 7812, etc from multiple vendors) of voltage regulators that have been available since 1970.

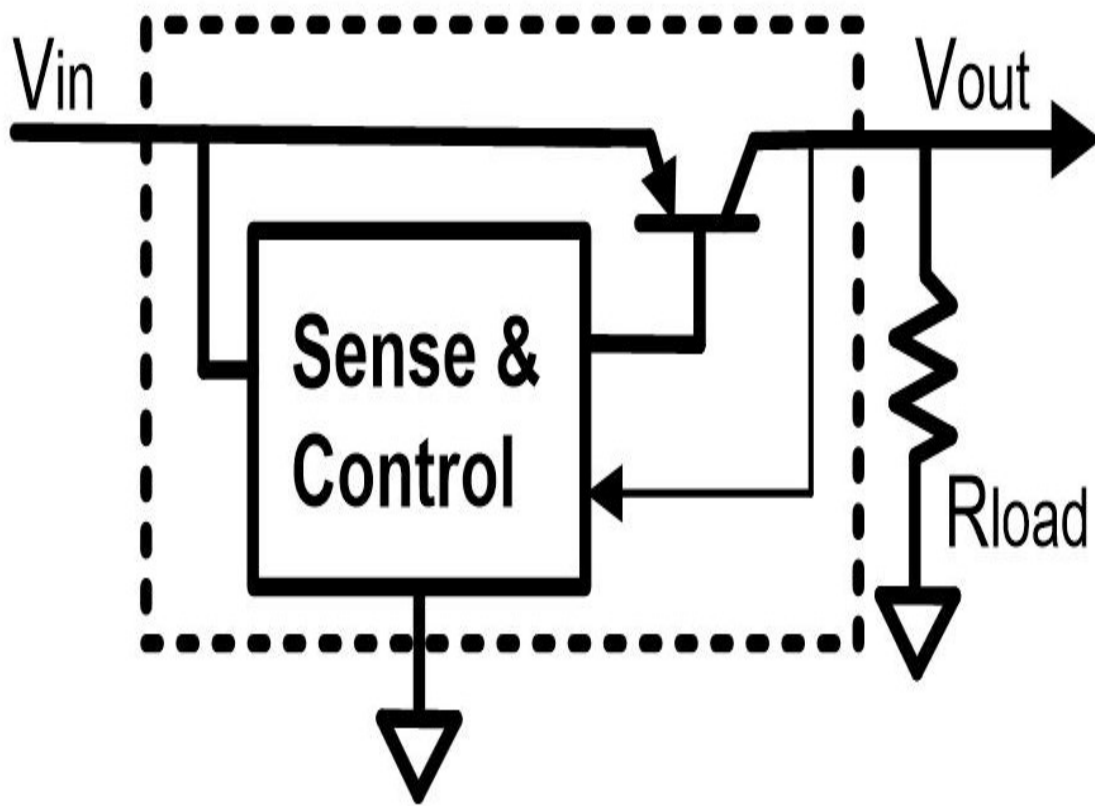


Figure 4-11. Linear Voltage Regulator, Low Drop Out

A more modern variant of the linear regulator is the LDO (Low Drop Out) device (Figure 4-?) which uses a power transistor that requires the control voltage presented to the power transistor to go down to increase the output voltage. This is done with either a PNP bipolar transistor or a PMOS MOSFET. By doing this, the necessary input voltage to maintain a proper output is greatly reduced. LDO devices typically will work with an input voltage that is 100mV greater than the output. Due to that capability, the LDO is widely accepted as a more versatile linear regulator. The LDO still has the limitations of being an active voltage divider method, and the

associated inefficient use of power. LDO regulators can also exhibit some feedback stability issues if not properly loaded and compensated, as per the vendor's recommendations for their specific devices. Read the vendor's application documentation to avoid problems.

Switching Stepdown (Buck) Converter

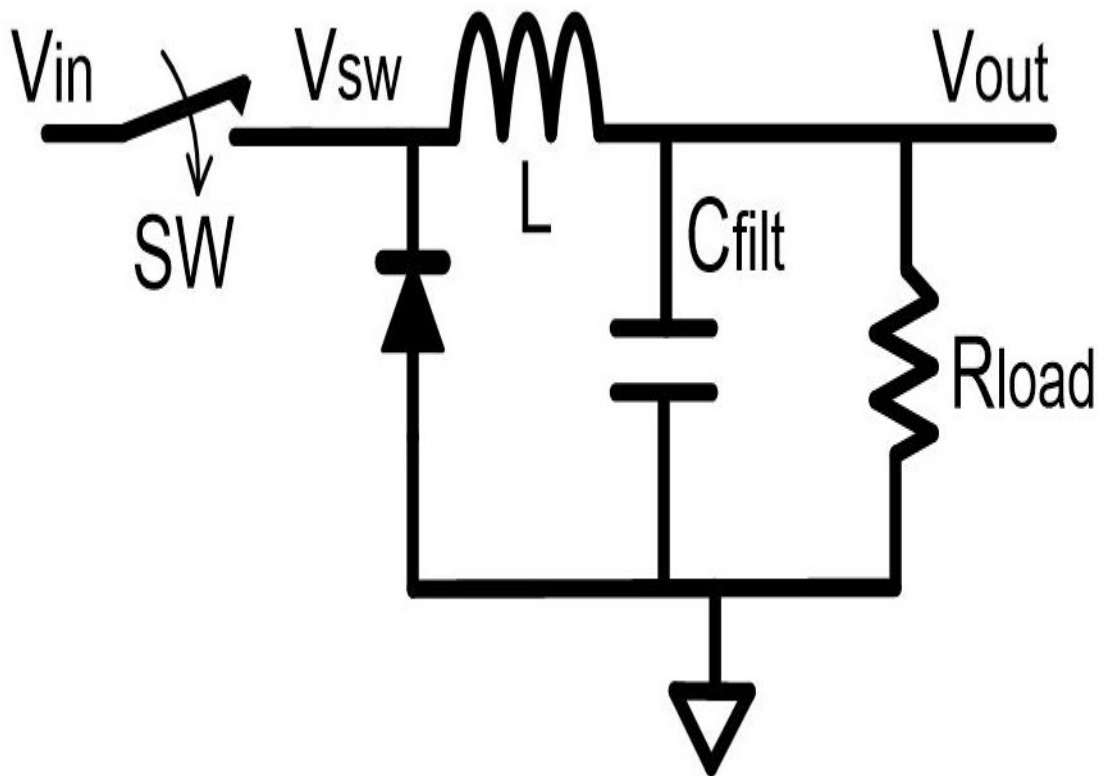


Figure 4-12. Switching Stepdown Converter Conceptual

The switching stepdown converter, commonly called a “Buck converter” uses varying duty cycle methods to reduce an input voltage (Figure 4-?). The switch is turned on and off at a high frequency (10 KHz to 1 MHz are common) while a control feedback system determines the time period that the switch remains on. A higher percentage of on time yields a higher output voltage (Figure 4-?).

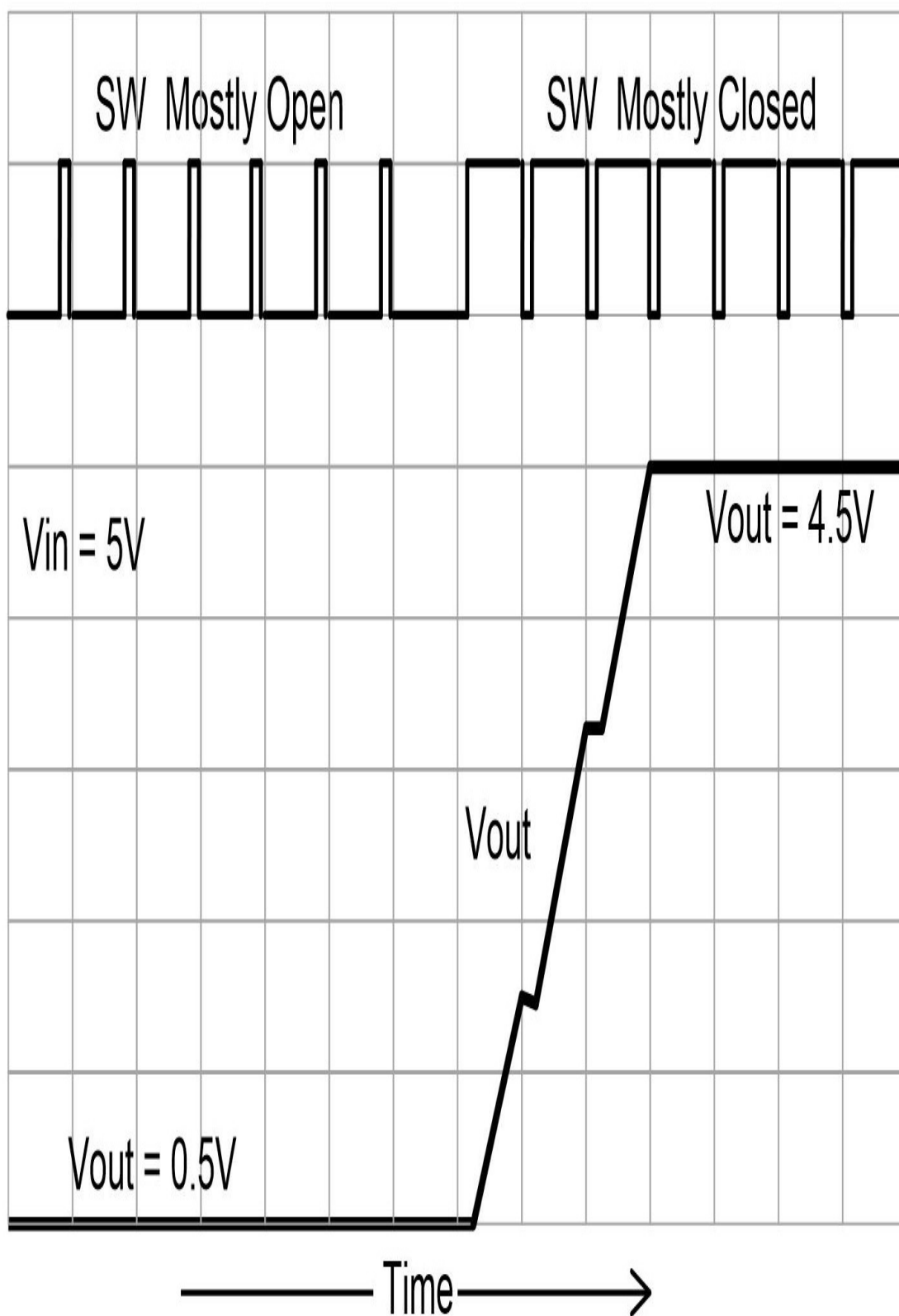


Figure 4-13. Switching Constant Frequency Variable Duty Cycle

When on, near 100% duty cycle, the output voltage will be close to the input voltage, a reduction in duty cycle will provide proportionally lower voltages at the output. Energy storage for a steady output voltage is maintained on the capacitor, which is fed through the inductor by either the V_{in} (switch closed) or the magnetic field collapse of the inductor (switch open). The diode serves as a passive switch, limiting the input side of the inductor to not transition below ground when the switch is open. Higher switching frequencies, allows smaller value inductors and capacitors, thus providing small size and low cost implementations. The efficiency of the Buck converter is very high, typically over 90%, with some losses due to non ideal component characteristics.

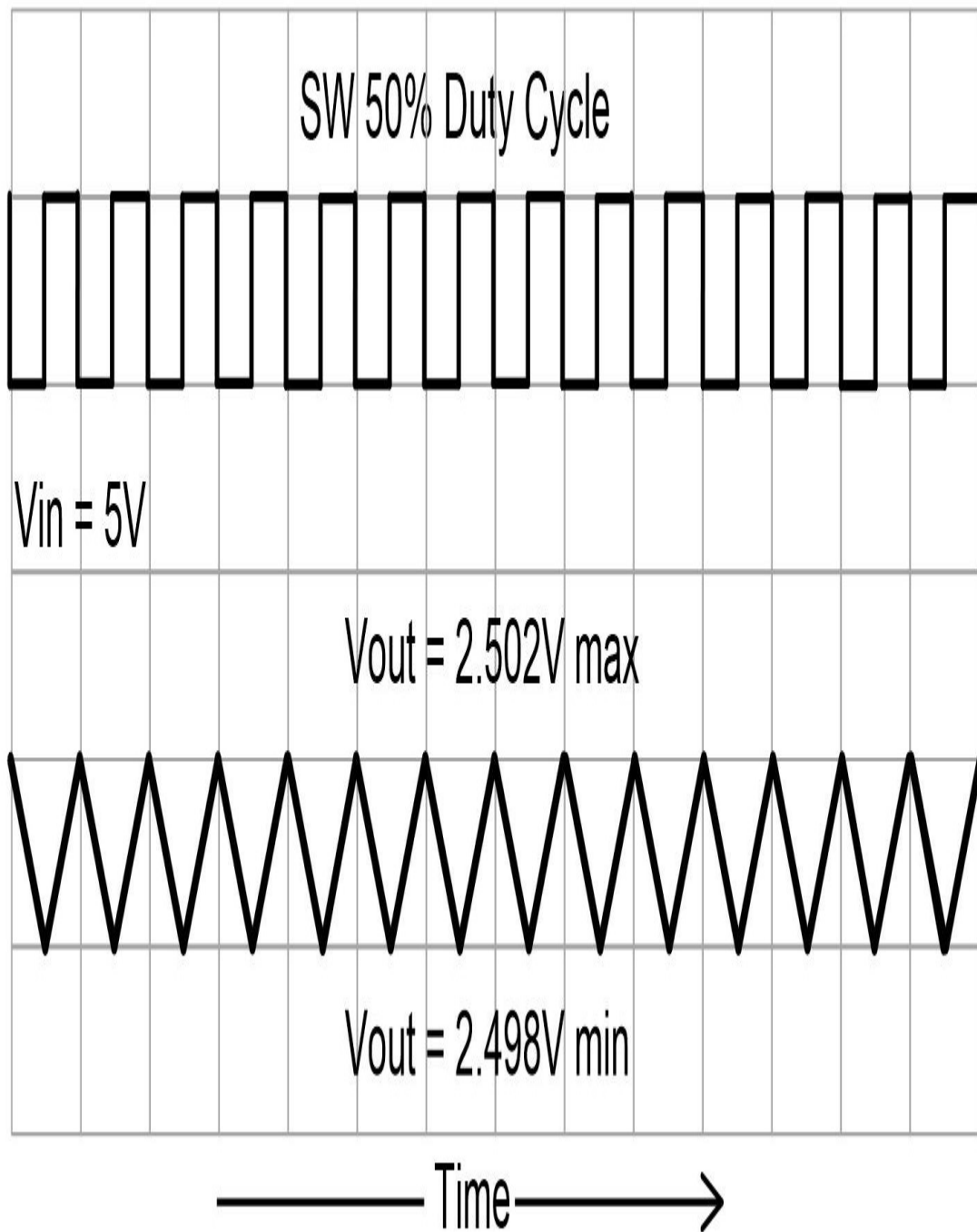


Figure 4-14. Switching Converter Noise on Output

Although the overall output voltage remains constant under load, close inspection of the output voltage shows that the switching creates some

variance in the output (Figure 4-?). All switching regulators will exhibit this behavior, a variant of a few millivolts is common in steady state situations, with the amplitude dependent on the switching frequency, output current load, inductor value, and capacitor value. Because of this switching noise, the use of switching regulators might not be appropriate in some noise sensitive situations. However, this power supply noise is still suitably low to provide acceptable power quality for most applications. Switching regulators are a high efficiency and versatile solution for all digital power situations, and can be used for some analog circuits. Use of switching regulators to power analog circuits needs to be carefully considered on a case by case basis.

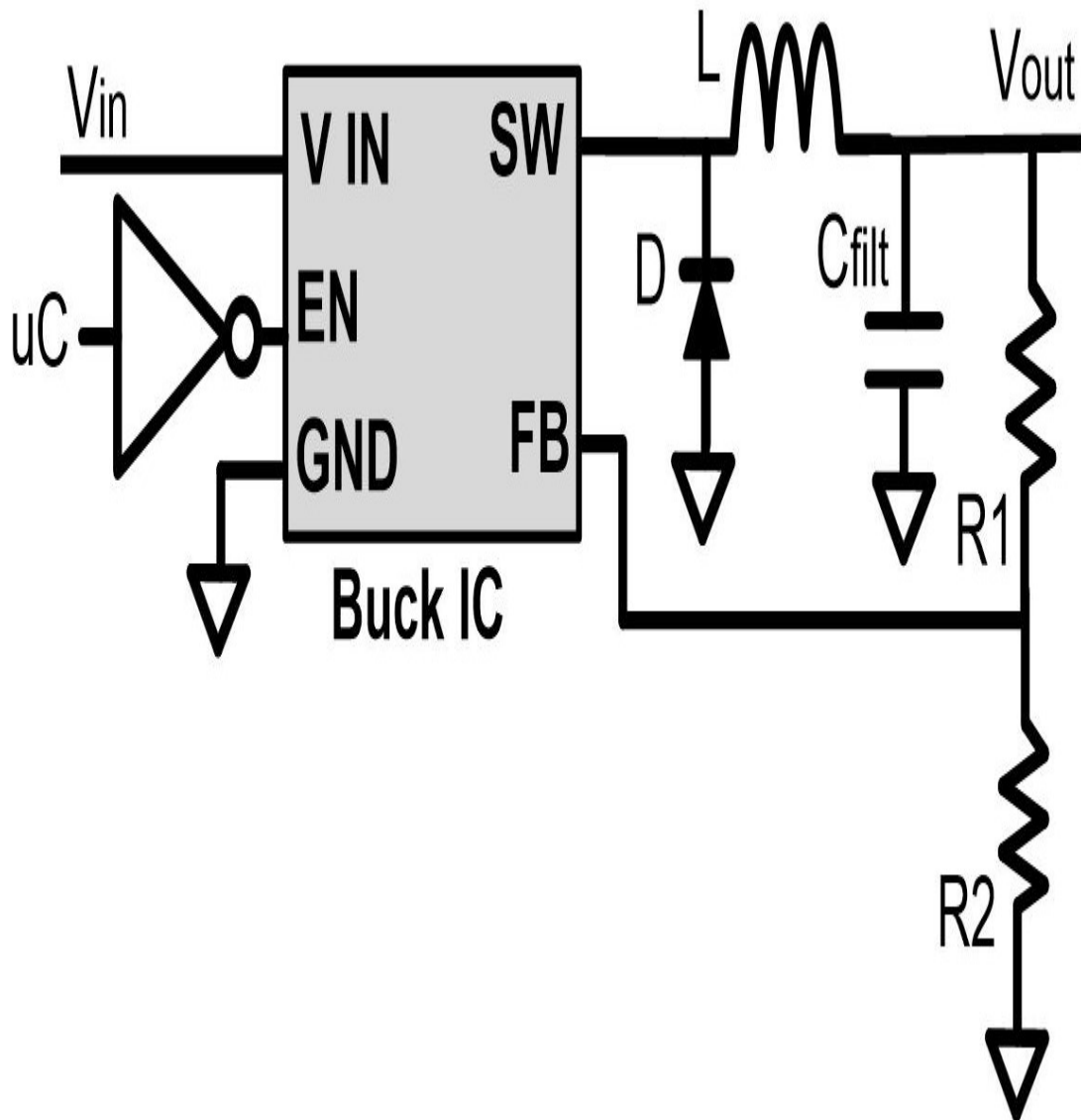


Figure 4-15. Switching Stepdown Converter Implemented

A typical implementation is shown, with $R1$, $R2$ allowing a selectable output voltage via feedback (FB) sensing, and the capability to enable (EN) the device by an external digital controller.

Implementations of the Buck converter may include variation in switching frequency, the diode may be internal to the IC, high power versions may use an external power transistor switch, an active switch can be used in place of

the diode for better efficiency, and the feedback sense resistors may be internal to the IC for fixed voltage versions.

Switching Step Up (Boost) Converter

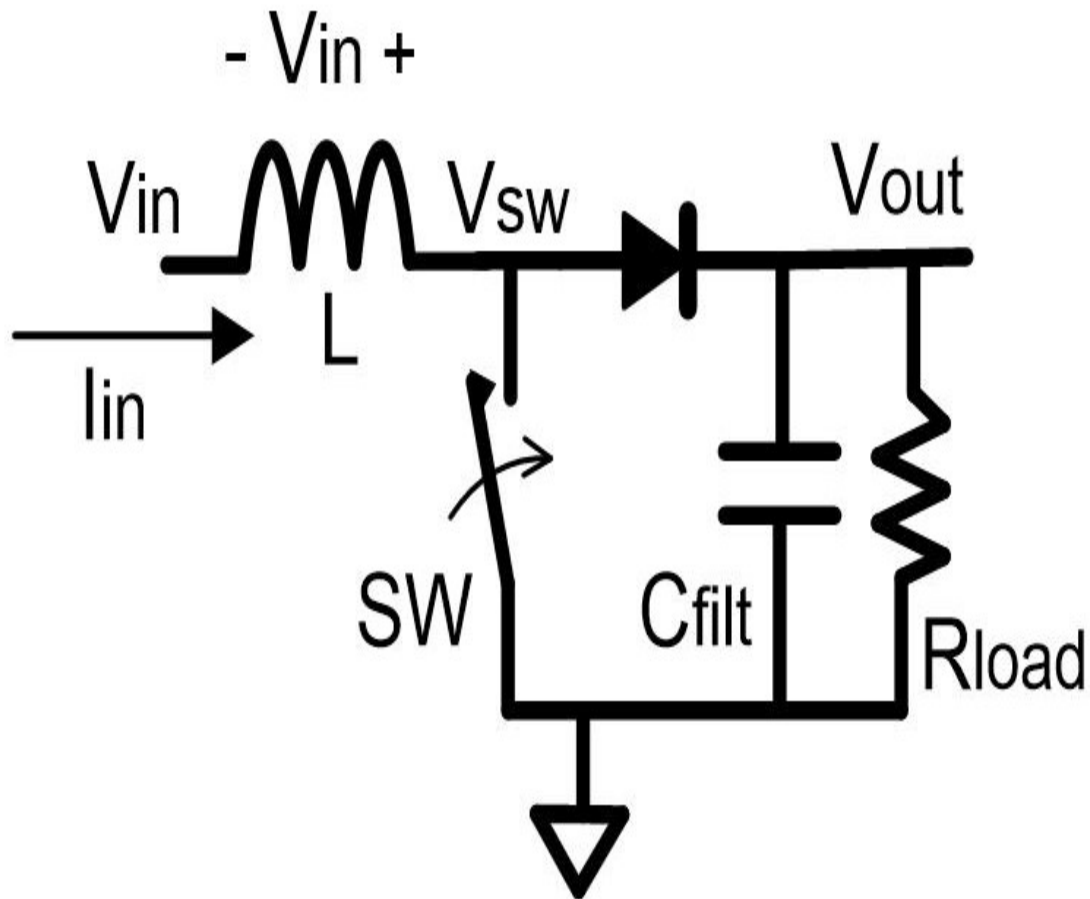


Figure 4-16. Switching Boost Converter, Conceptual

The switching step up converter, known as a “Boost converter” uses the reactive characteristics of an inductor to create a voltage above the input voltage (Figure 4-?). Closing the switch causes the current to ramp up in the inductor. Opening the switch causes the $V = L \, di/dt$ to create a positive inductor voltage at the V_{sw} node, which is in series with the V_{in} voltage. This forward biases the diode, dumping charge into the C_{filt} , resulting in a V_{out} above the input voltage V_{in} . The switch is turned on and off at a high frequency (10KHz to 1MHz are common) while a control feedback system

determines the duty cycle of the switch. A higher percentage of the open switch dumps more current into C-filt resulting in a higher output voltage. Energy storage for a steady output voltage is maintained on the capacitor, which is fed through the diode when the switch is open. The diode also serves as a passive switch when the switch is closed, isolating the C-filt while the magnetic field in the inductor is replenished. Similar to the Buck converter, the use of high switching frequencies allows the use of smaller value inductors and capacitors enabling small size and low cost implementations.

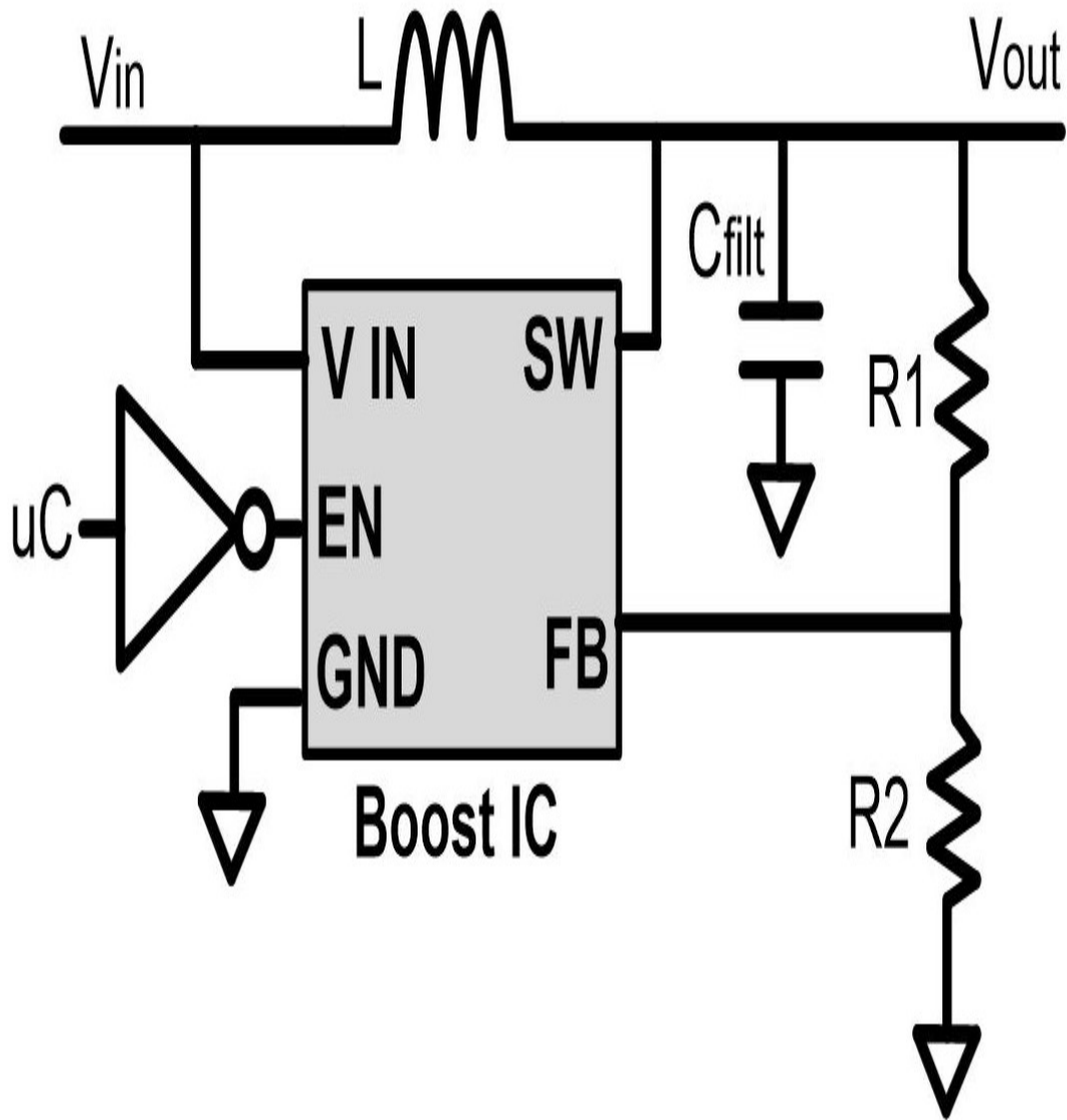


Figure 4-17. Switching Boost Converter, Implemented

A typical implementation is shown with $R1$, $R2$ allowing a selectable output voltage via feedback (FB) sensing, and the capability to enable (EN) the device by an external digital controller.

Variations on the Boost converter may include changes to switching frequency, the diode may be internal to the IC, high power versions may use an external power transistor switch, an active switch can be used in place of

the diode for better efficiency, and the feedback sense resistors may be internal to the IC in fixed voltage variants.

Switching Buck-Boost Converter

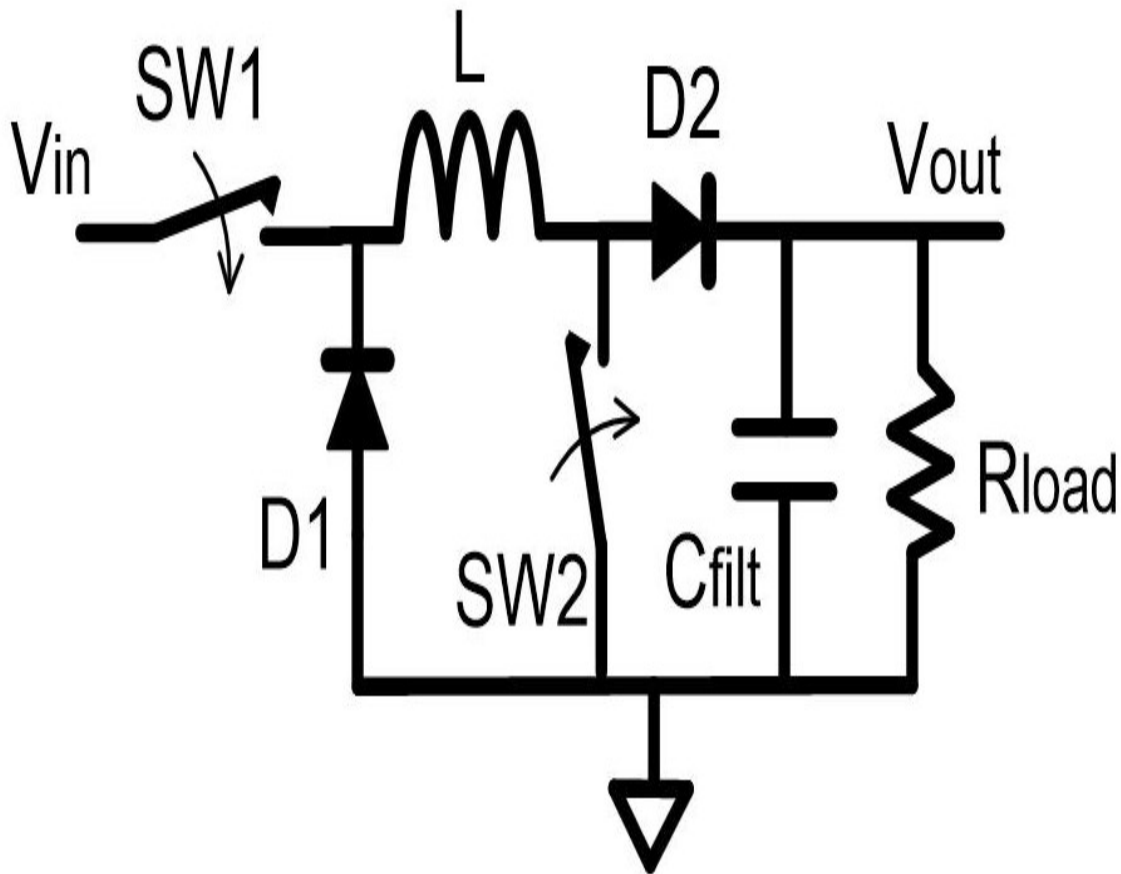


Figure 4-18. Switching Buck and Boost Converter, Conceptual

The Buck-Boost Converter can both raise and reduce the output voltage. When $SW1$ is actively controlled and $SW2$ is left open, the device functions as a Buck converter. With $SW1$ closed and $SW2$ actively controlled, the device functions as a Boost converter. In a typical application (Figure 4-?) the internal control circuitry determines which mode is needed based upon input voltage available and the target output voltage. The combined Buck-Boost converter is a useful tool when a widely variable input power supply

is used, such as a battery. In these applications, the converter reduces the battery voltage to the target output voltage until the battery has discharged enough to drop below the output target and then changes mode to a Boost converter to extend the useful battery range.

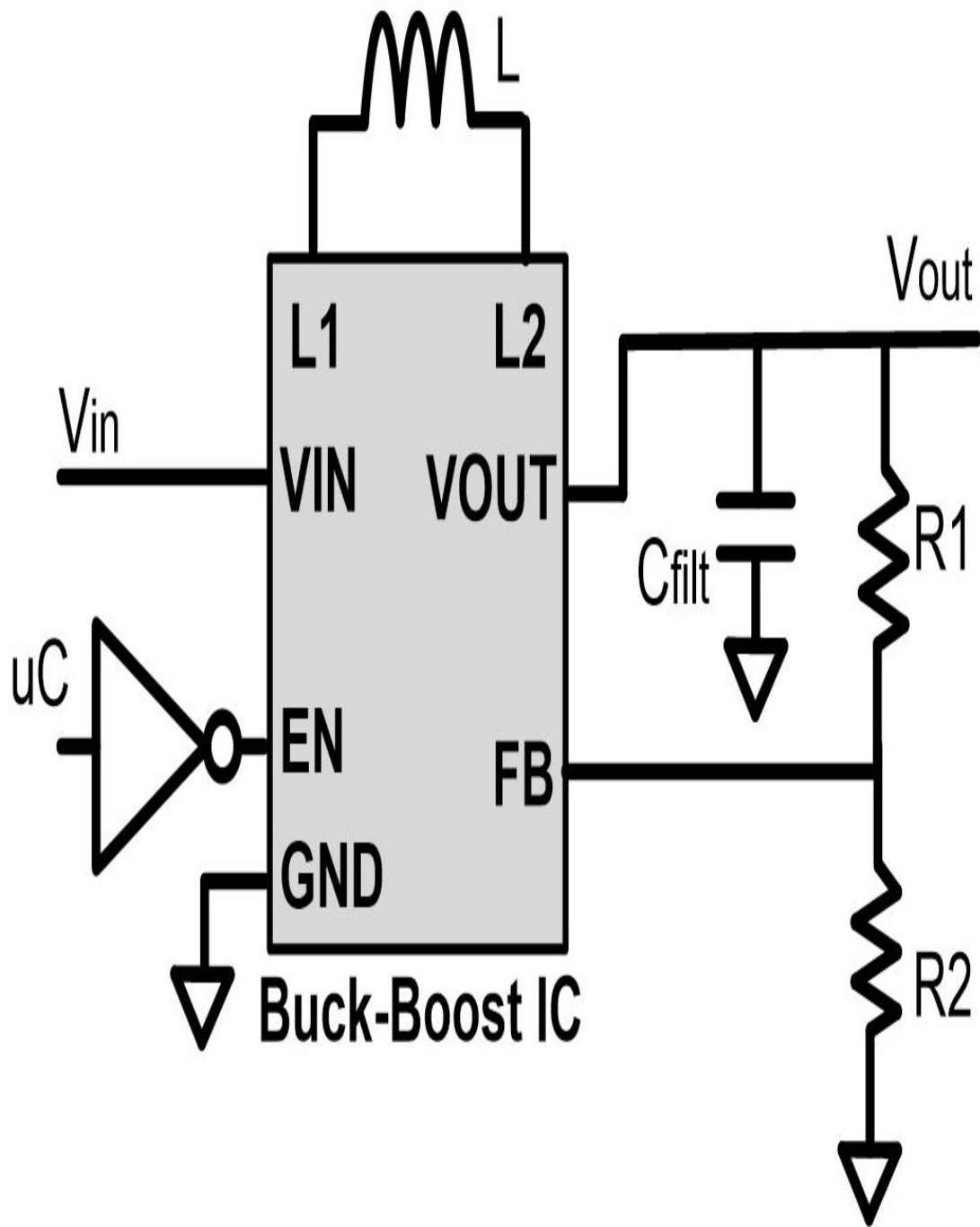


Figure 4-19. Switching Buck and Boost Converter, Implemented

Picking Regulators and Configuring Power System

Putting a power system together is best illustrated with some case studies. The first example has a 12V motor driver system, a block of logic that functions from 2.5V, the host microcontroller needs 3.3V and some analog circuitry that needs 5V. This is a plug in device that uses an AC/DC converter.

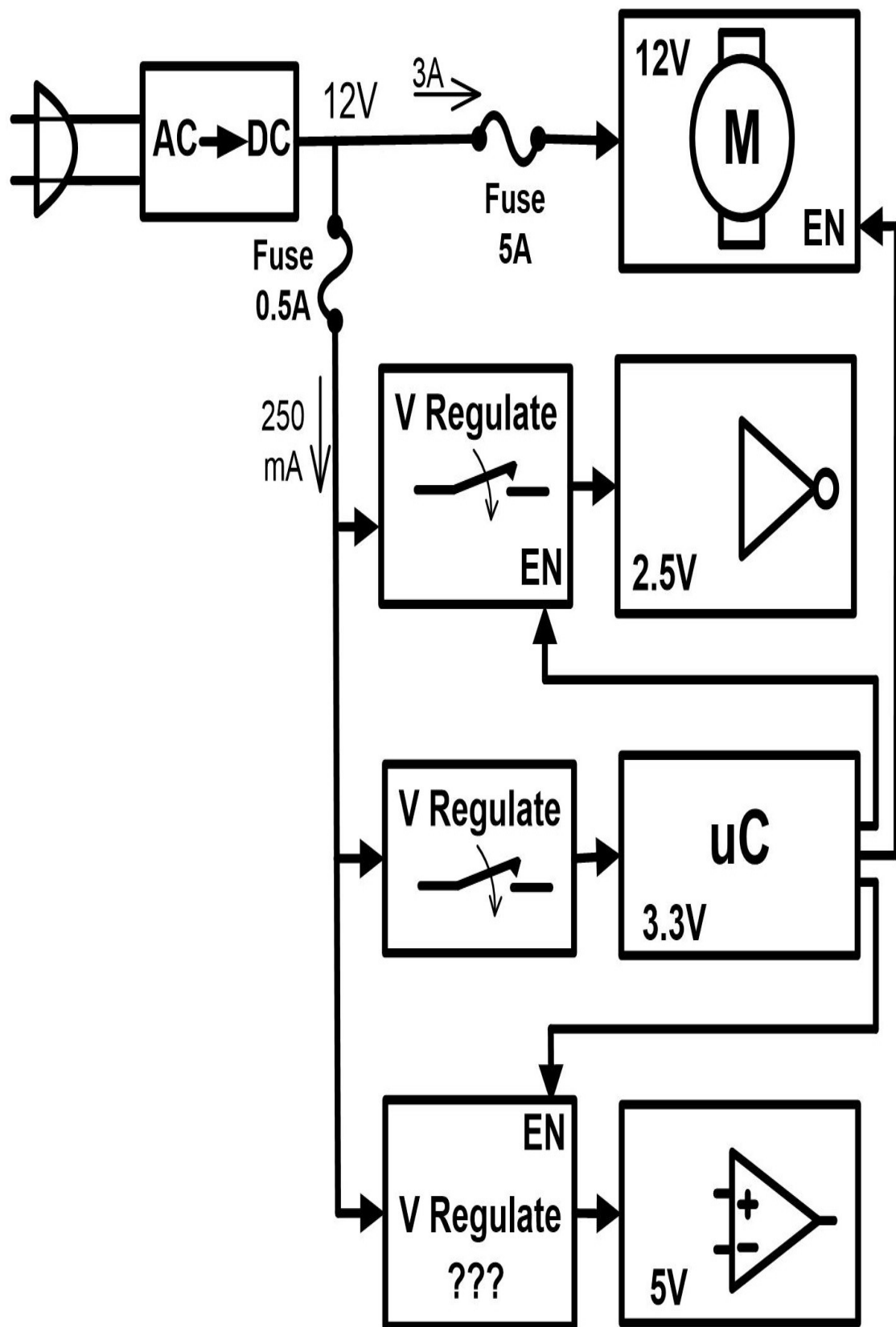


Figure 4-20. Case Study Multiple Power Supplies Plug in Device

Configuring the power here is mostly straightforward. A 12V output offline switcher, does the AC/DC conversion and will be used to directly power the motor driver. Motors can be high current, but a limited amount of voltage dip is tolerable when starting. Motors are also immune to most power noise and some power variance can be compensated for by the motor controller. The current rating of the AC/DC converter needs to support the start surge current of the motor, steady state motor current under maximum load and the other electronics. When selecting an AC/DC converter, in addition to voltage and current specifications, look for devices that are self protective and current limiting, and consider the output ripple and noise specifications carefully as your application may have special requirements. If your product will go to markets outside the USA, consider devices that will function off of both 50/60 Hz and 110/220 VAC without need to change settings or connections.

The 2.5V logic, can be readily powered from a Buck converter. The microcontroller needs 3.3V and can also use a Buck converter. Power for the microcontroller is always on, and as the central control of the system, it enables all other power supplies and other devices. This allows selective shutdown of peripheral devices, and allows an organized power up/down sequence, that provides predictable and safe system behavior.

The analog block that needs 5V power is not straightforward. More needs to be known, namely, how much current is needed and how critical is low noise power performance to the circuit? If the circuit can tolerate the power noise associated with a switching supply, it's the most energy efficient method. Considering this is a plug in the wall application energy efficiency is not a high priority. Using a linear regulator here gives a quiet supply but would be energy inefficient. However, a 5V regulator here has 12V input and that 7V drop across the linear regulator, will be dissipated as heat. In this case, if the current rises to 140 mA, the power burned in the linear regulator rises to 1W. If the current is low, and low noise power is important, a linear regulator would work. Check power dissipation in any

linear regulator carefully. In many cases the switching regulator with some damped filtering may be a viable alternative.

Finally, the system uses fuse protection, of 5A on the motor circuit which draws a worst case current of 3A, and a 0.5A fuse on the electronics power. Using two fuses is due to the high current motor circuitry. If a single fuse were used, the low current electronics would have to exceed 5A and that would probably be beyond the destructive current level of the electronics. Other possible options here include a PTC on the electronics for protection, and a circuit breaker on the motor circuit, if the motor application is in an environment that stalls or goes into overcurrent loading frequently.

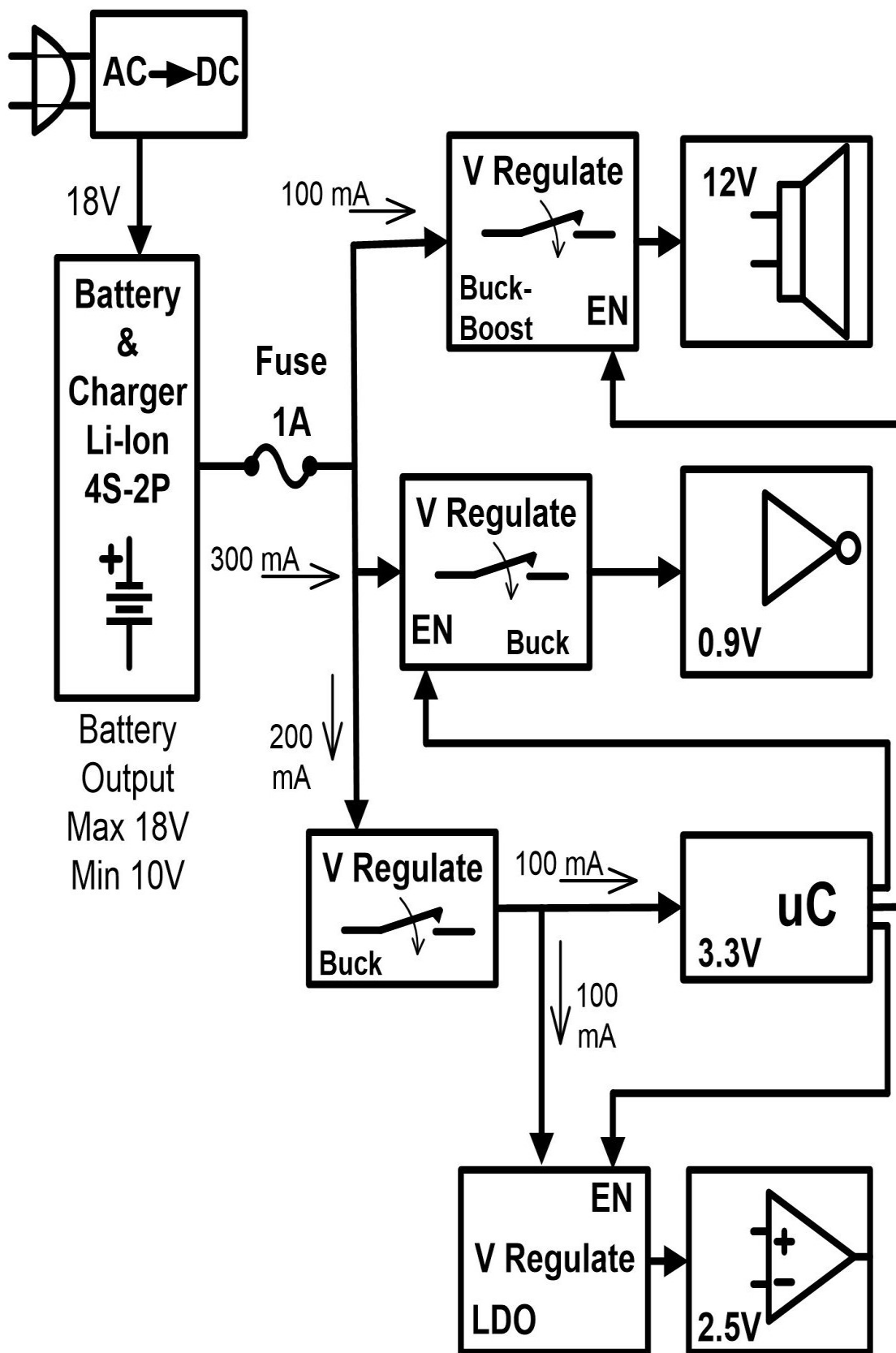


Figure 4-21. Case Study Multiple Power Supplies Battery and Charger

Looking at Figure 4-?, the power system includes a battery and charger system. This example uses a Lithium battery pack with 4 batteries in series. When the charger is applied, the power system sees 18V, and when the batteries are largely discharged the power system sees 10V. Everything needs to function at both voltage extremes. One circuit needs 12V, which requires a Buck-Boost converter, that functions in Buck mode until the battery pack goes below 12V and then changes to Boost mode. Examine the 12V circuit for noise requirements which may indicate a need for a filter at the converter output. The digital block that uses 0.9V power is readily powered from a Buck converter. The microcontroller runs off 3.3V and is always powered on from the battery system, while controlling the power supplies to the rest of the system. The final block runs off 2.5V and requires low noise power input. This device is a good candidate for a linear regulator, using an LDO that is powered off of 3.3V. The 0.8V of headroom is sufficient for a LDO device to function without a large voltage drop and the associated power lost as heat. Carefully examine this noise sensitivity of the circuitry and consider the need for filtering both before and after the LDO to provide a quiet 2.5V. High frequency noise can couple across voltage regulators and a strategy of filtering input and output may be needed in very noise sensitive situations.

Both rechargeable battery systems and disposable batteries, will have similar needs to remain functional over a wide variance in the supply voltage.

Including Monitors for the Power Supplies

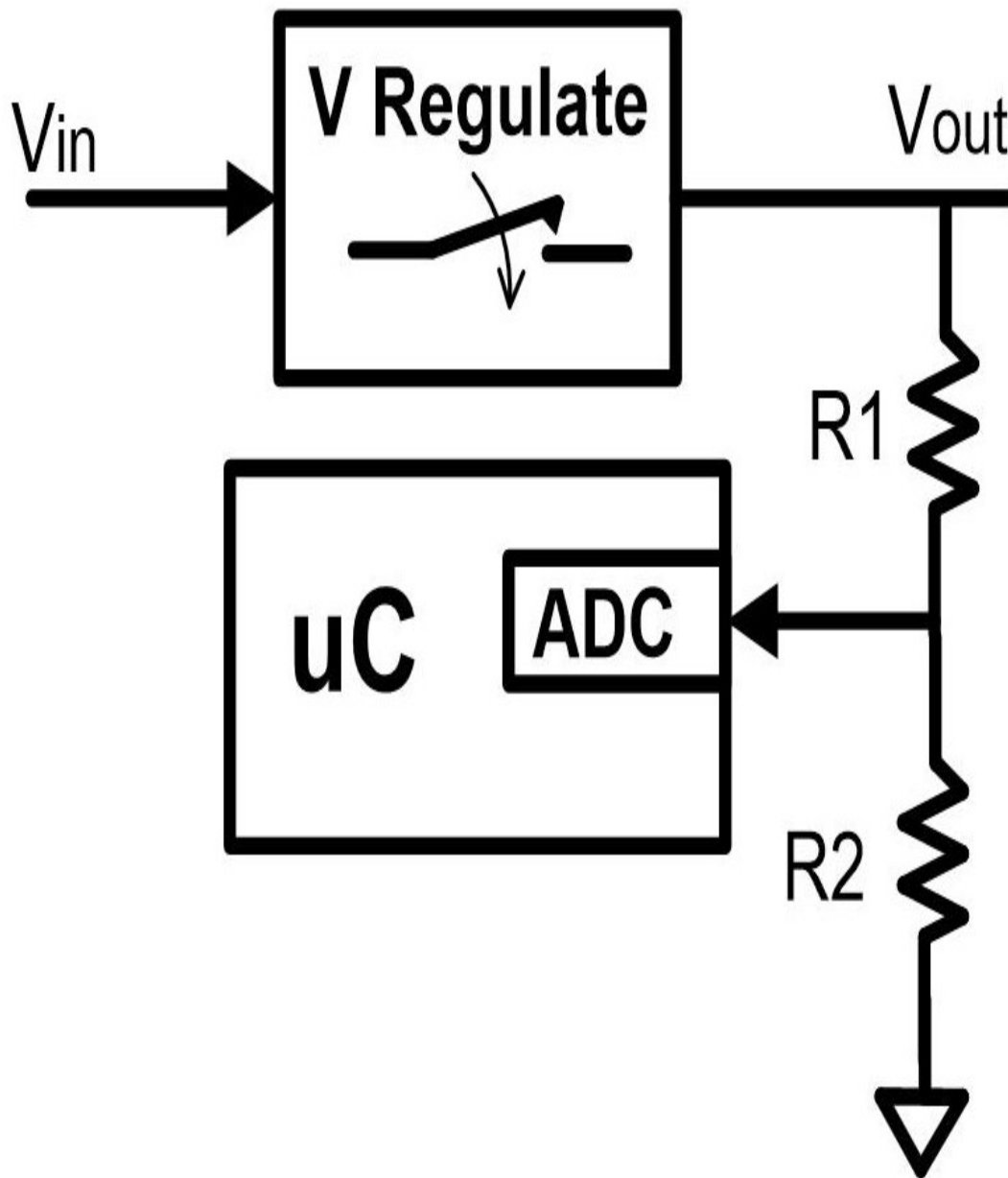


Figure 4-22. Monitoring Power Supplies

A useful addition to the power control system is the capability to monitor the voltages out of the regulators. Adding a voltage divider to an output voltage (Figure 4-?) and feeding that back to an ADC port available on the host microcontroller can be useful in checking power sources, or assisting

in a sequenced power on-off cycle. The voltage divider exists to scale V_{out} down to the mid scale range of the ADC in the microcontroller.

Bypass, Decoupling and Filtering Techniques

Many of the multipurpose microcontrollers on the market have some limited analog capability, such as ADC ports. Some of these controllers separate the power with both digital and analog power connections. This provides a good opportunity to clarify the difference between a bypass capacitor and a decoupled power connection.

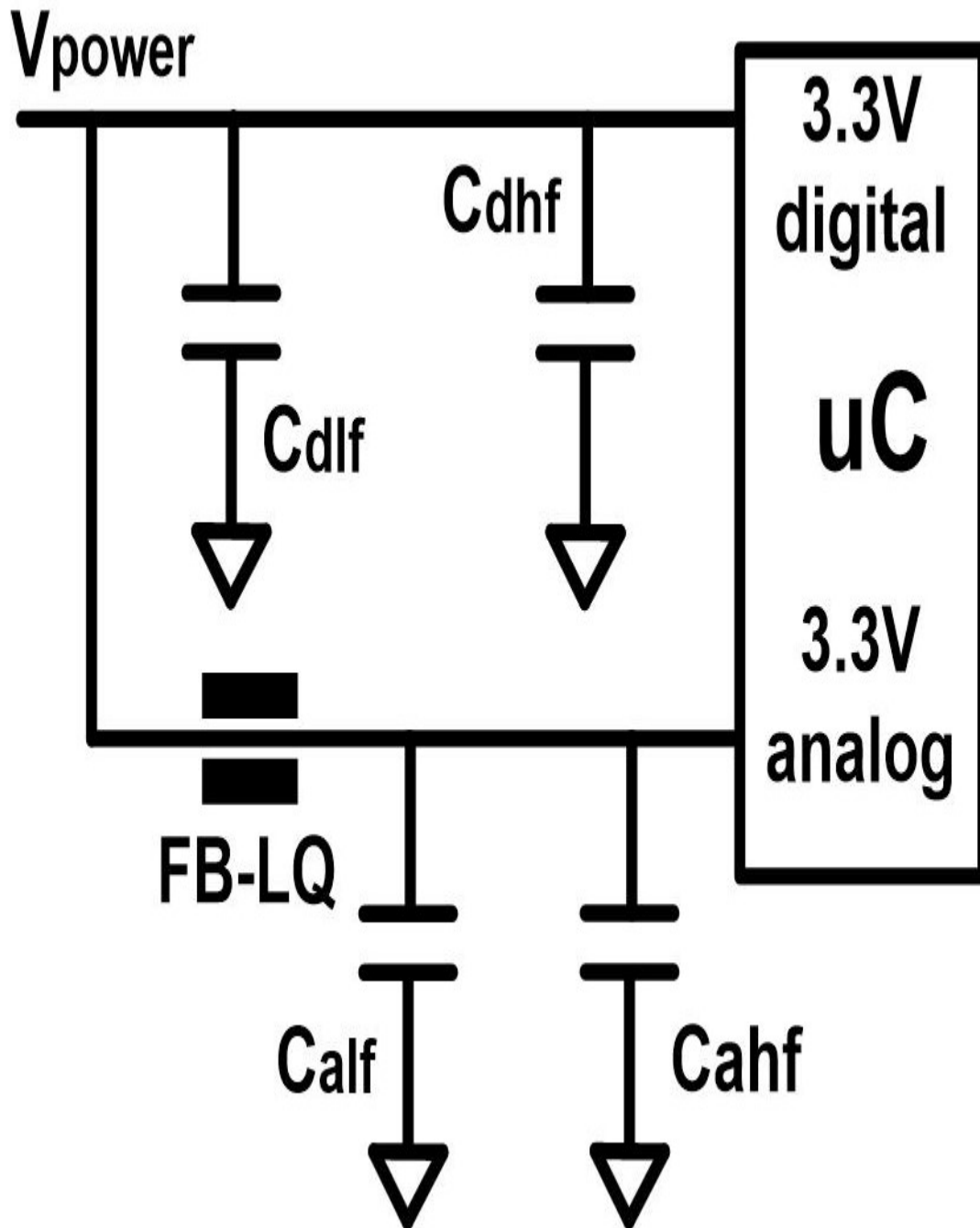


Figure 4-23. Microcontroller Power, Bypass Digital, Decoupled Analog

Frequently you can decouple the analog supply pins locally, using a ferrite bead and then both lower frequency bypass (C_{alf}) and high frequency bypass (C_{ahf}) directly at the pin. A damped low pass power filter can be used in place of the ferrite bead if greater rejection of power input noise is needed. Figure 4-? illustrates the difference between bypassing and

decoupling of power connections. The two terms are often interchanged or used incorrectly. Here, the capacitors (C_{dlf} , C_{dhf}) bypass the V_{power} to maintain high frequency stability, while the combination of the ferrite bead (FB-LQ) and capacitors (C_{alf} , C_{ahf}) serve as a low pass filter to decouple the analog power pin from the digital switching noise present on the V_{power} node. Decoupling involves series impedance in the power connection, and can be useful in limited situations where the decoupled load does not need heavy transient currents.

Radiated Noise Reduction, using RC Snubbers, Ferrites, Filters

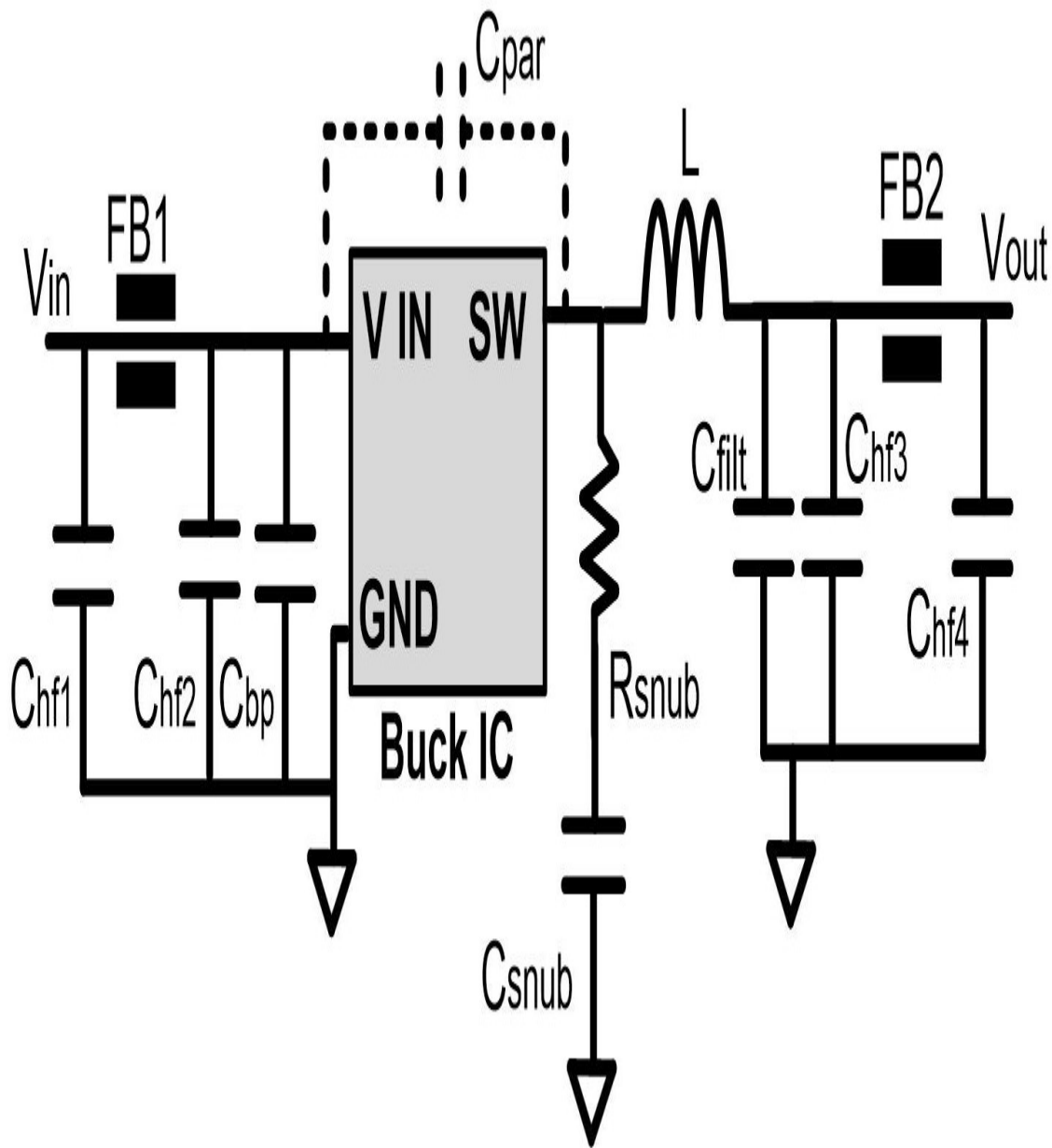


Figure 4-24. Noise Reduction Using Snubbers and Ferrites

In addition to switching regulators putting out somewhat noisy power, they will also inject noise back into the power grid that they are fed from. Implementation of many switching converters involve components (Figure 4-?) to reduce the high frequency noise that they pump onto the power grids and radiating EMI. In this example, both input and output are lowpass

filtered with a ferrite and suitably sized high frequency capacitors on both sides of the ferrite. If more isolation is required, an RLC network damped lowpass network can be substituted. In addition to the input-output filtering, an RC snubber (R_{snub} , C_{snub}) is placed at the switched node. The RC snubber provides a low impedance grounding path for high frequency noise generated at this switched node. (FDI: EMI-ESD)

Specific values for the RC snubber are dependent on converter frequency and internal transistor sizes. The manufacturer of a specific switching converter should provide suitable component values for their devices.

Power Output Noise Reduction - Damped LPF Networks, Cascaded Regulators

Higher order filtering may be needed in some situations. For an in depth design approach to damped low pass filtering methods, the Tompasett paper (see further reading references) outlines a detailed design methodology. Other approaches here include using a switching converter which is followed by a linear regulator that reduces noise further. This method allows a low noise output, while keeping the power losses associated with linear regulators minimized.

Power Grid Current Surges Due to Digital Logic

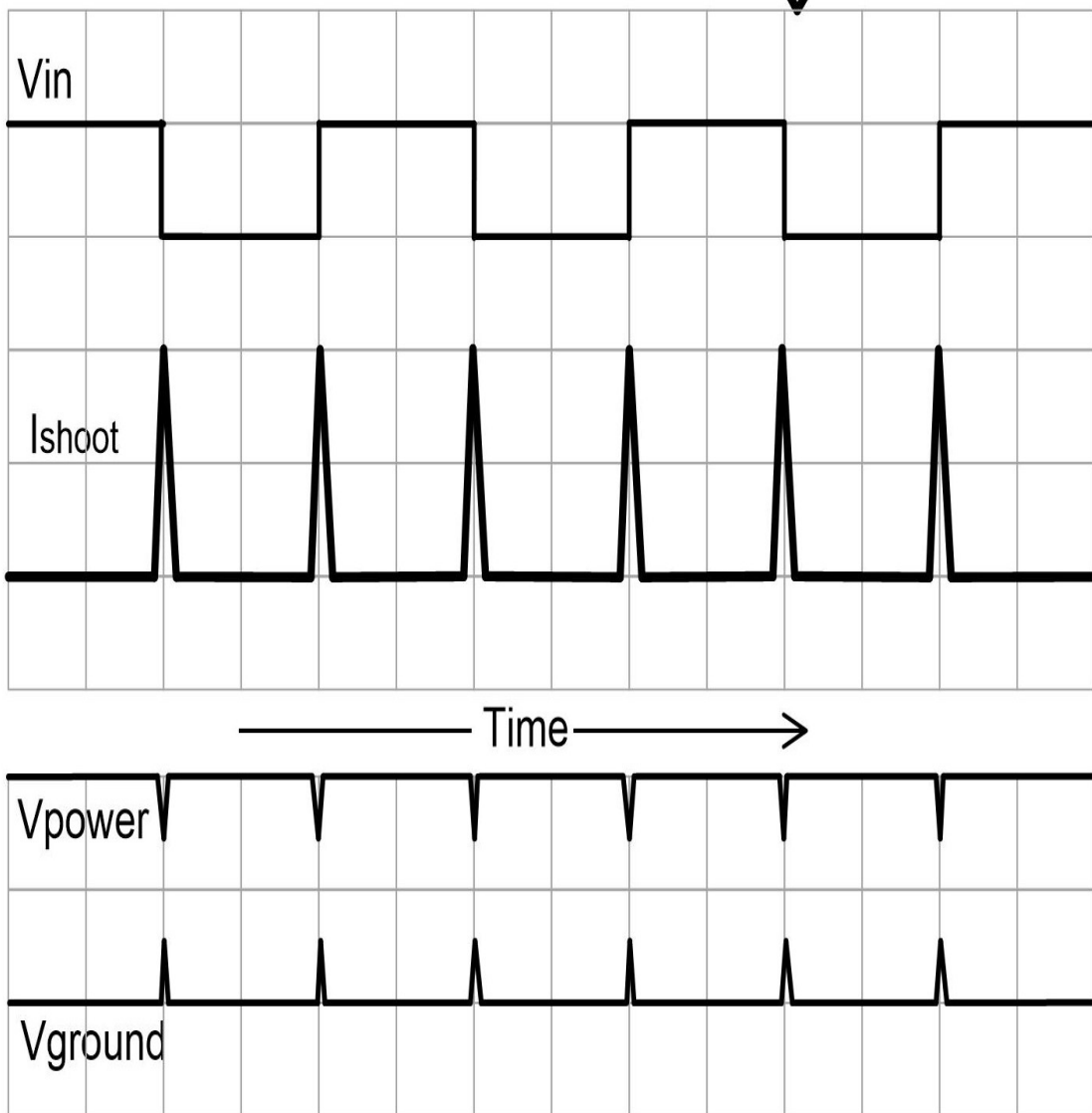
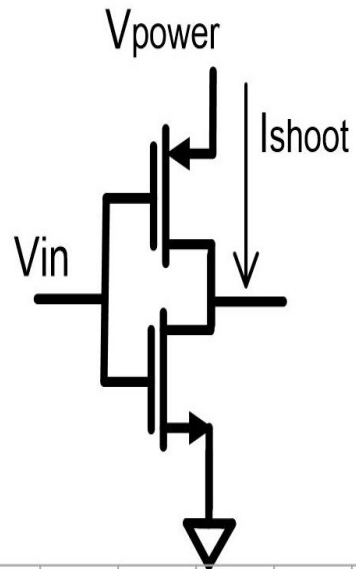
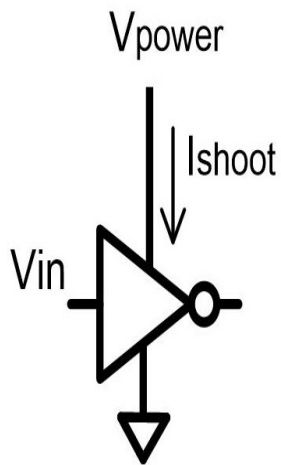


Figure 4-25. Digital Current Shoot Through

The use of digital CMOS logic poses special challenges in keeping the DC power grid stable. The transition switching of any CMOS logic causes both NMOS and PMOS transistors to both be on at the same time during the state transition. The result is a resistive short circuit between power and ground on the state transition of every logic gate. This current spike, known as shoot-through, replicated among the countless logic gates within any digital device, creates both wideband EMI and high frequency current demands on the power. This is especially problematic due to Input/Output drivers inside digital IC's due to the large transistors used within the chip. The noise spectrum is distributed in frequency and is not just the fundamental clocking frequency, but includes many harmonics as well. Any impedance in the power or ground paths creates variance in the local power and ground voltages.

Low Impedance Power and Ground Planes

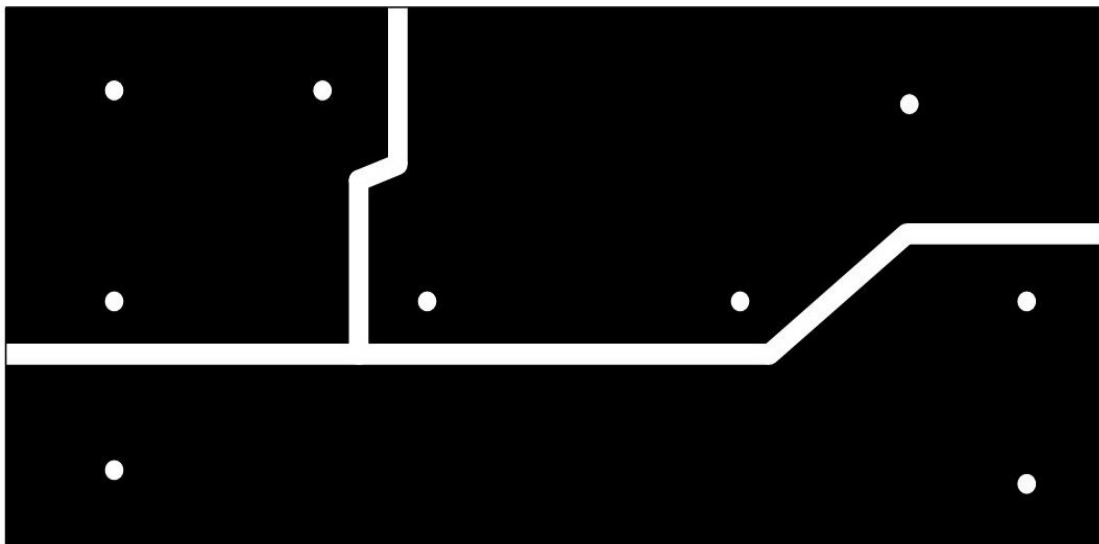
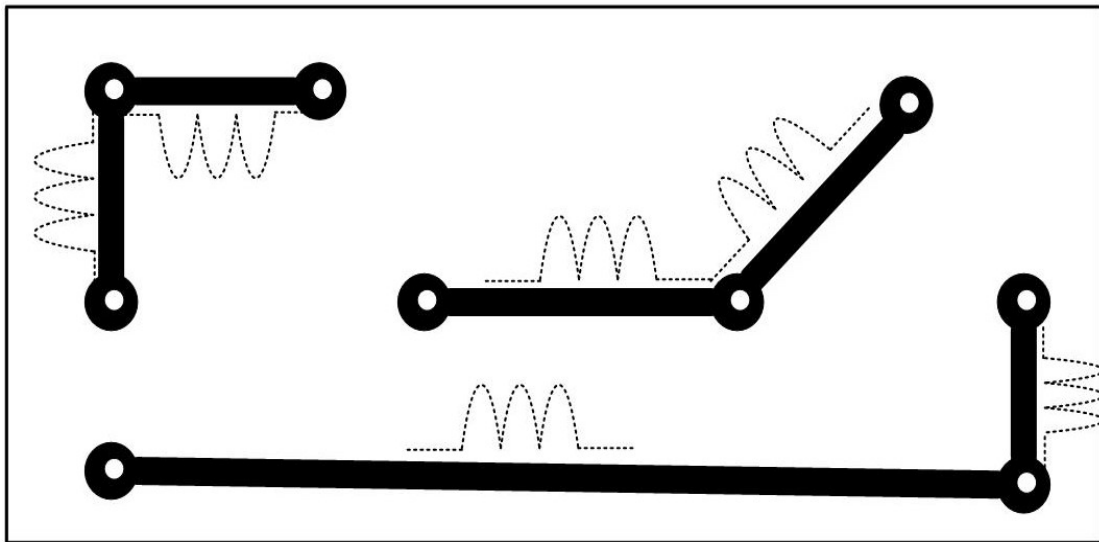
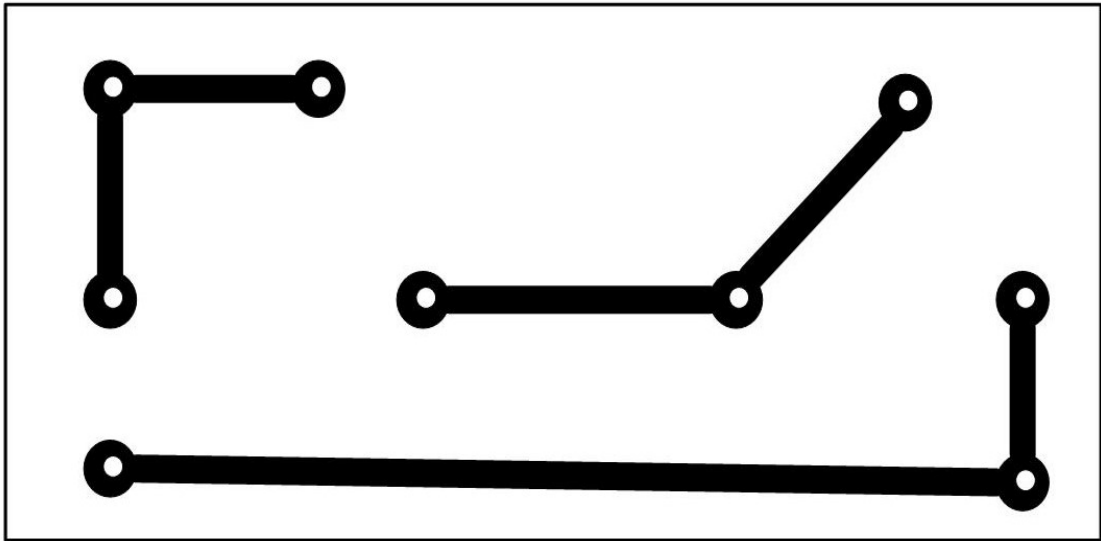


Figure 4-26. Power Connections With Low Inductance Interconnect

Due to current surges, the impedance of power and ground connections need to be kept as low as possible. For that, dedicated power and ground layers within the PCB are a must have. Point to point interconnects will not suffice. Figure 4-? illustrates a hypothetical power interconnect with three different power values, using point to point connections. All connections exhibit inductance and will cause the voltage across the power connection to vary dynamically as a function of any transient currents created by the devices being powered. Figure 4-? illustrates connecting the same points with a lower connection impedance. Placing this segmented power plane over a solid ground plane allows the capability to wideband stabilize power supply voltages with bypass capacitors.

Power Supply Bypass Filtering, Distributed Stabilization

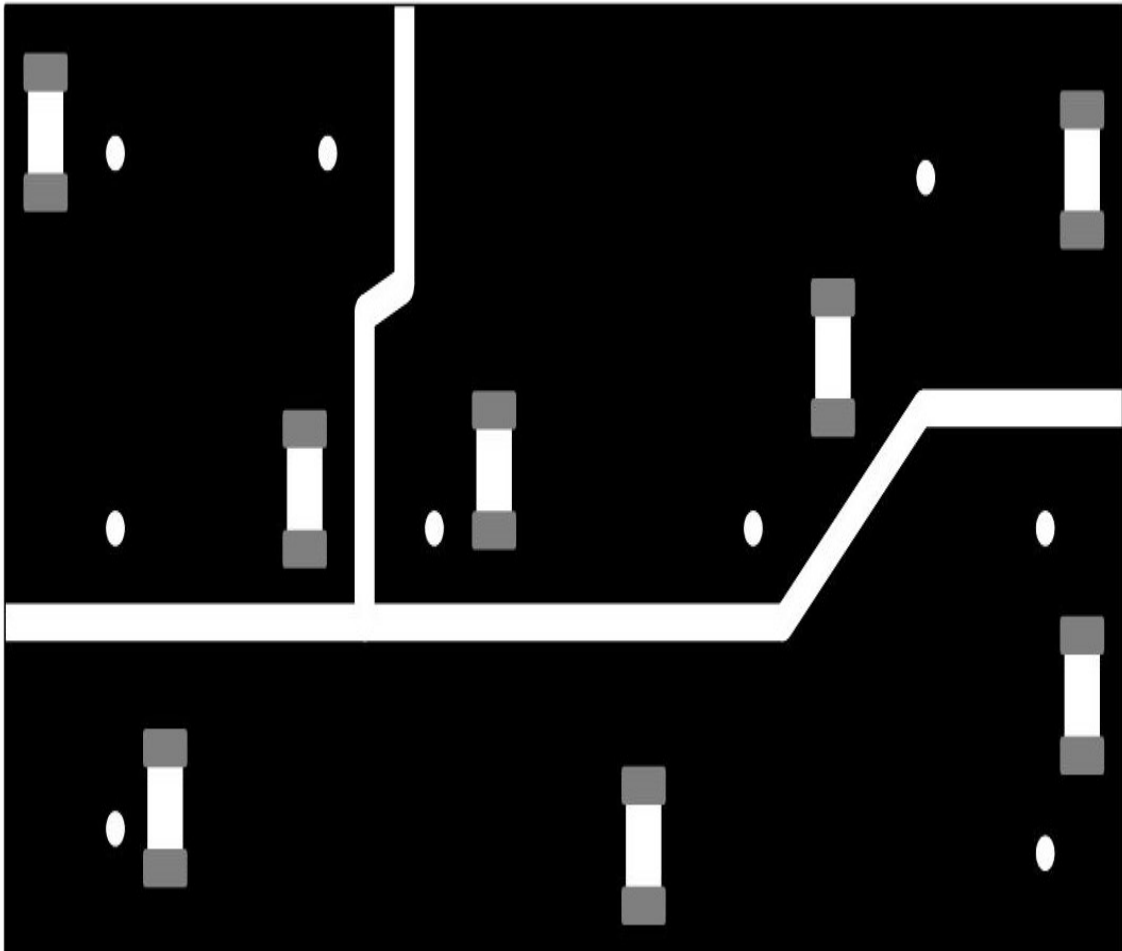


Figure 4-27. Distributed Stabilization of Power Grid

In addition to the low impedance interconnect strategy, the use of distributed bypass capacitance across the power plane (Figure 4-?) allows the high frequency stabilization of the power.

Typically, stabilization of a power grid has three parts:

- Bypass capacitors directly at all power pins on all IC's. (At chip bypass)
- Bypass capacitance distributed across the power plane. (Grid bypass)
- Bypass filtering at the voltage regulator output. (Regulator bypass)

A large PCB (i.e. a computer motherboard) with multiple high speed digital chips must have all of this. A small low frequency PCB (i.e. a handheld

remote for the TV) would be less demanding, and the grid bypass portion might be eliminated.

Bypass Capacitors at High Frequencies

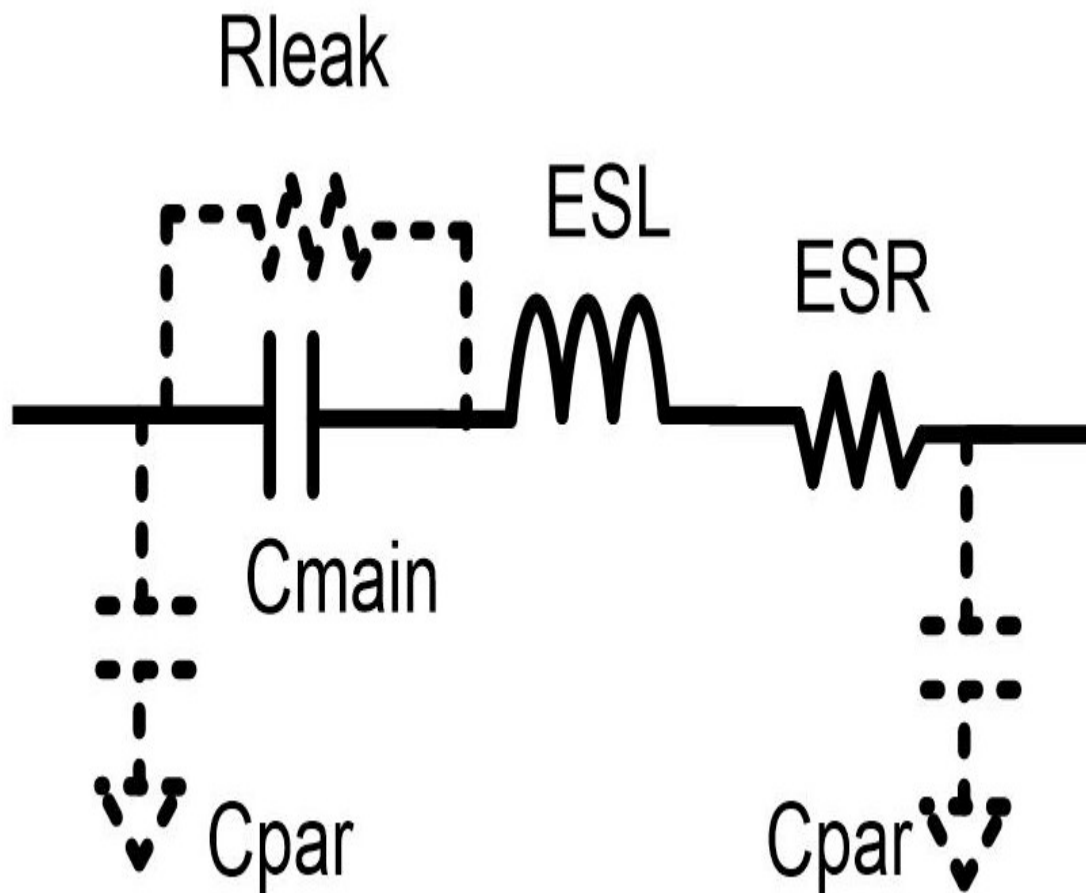


Figure 4-28. Equivalent Model for High Frequency Capacitors

The equivalent model for a surface mount “Multi Layer Ceramic Capacitor” (SMT MLCC) is shown in Figure 4-?. The elements of this are:

- C_{main} - principle body of the capacitor
- C_{par} - parasitic capacitance to the underlying PCB
- ESR - Equivalent Series Resistance of the distributed device resistance

- ESL - Equivalent Series Inductance of the device
- Rleak - leakage resistance between capacitance plates

The two important parameters for power bypass are ESR and ESL. The ESR limits the lowest impedance, and the ESL limits the maximum frequency where the device works effectively as a capacitor. These characteristics need to be considered in the selection of bypass capacitors.

1uF SMT 0603 Cap, Self Resonance @ 5.5 MHz

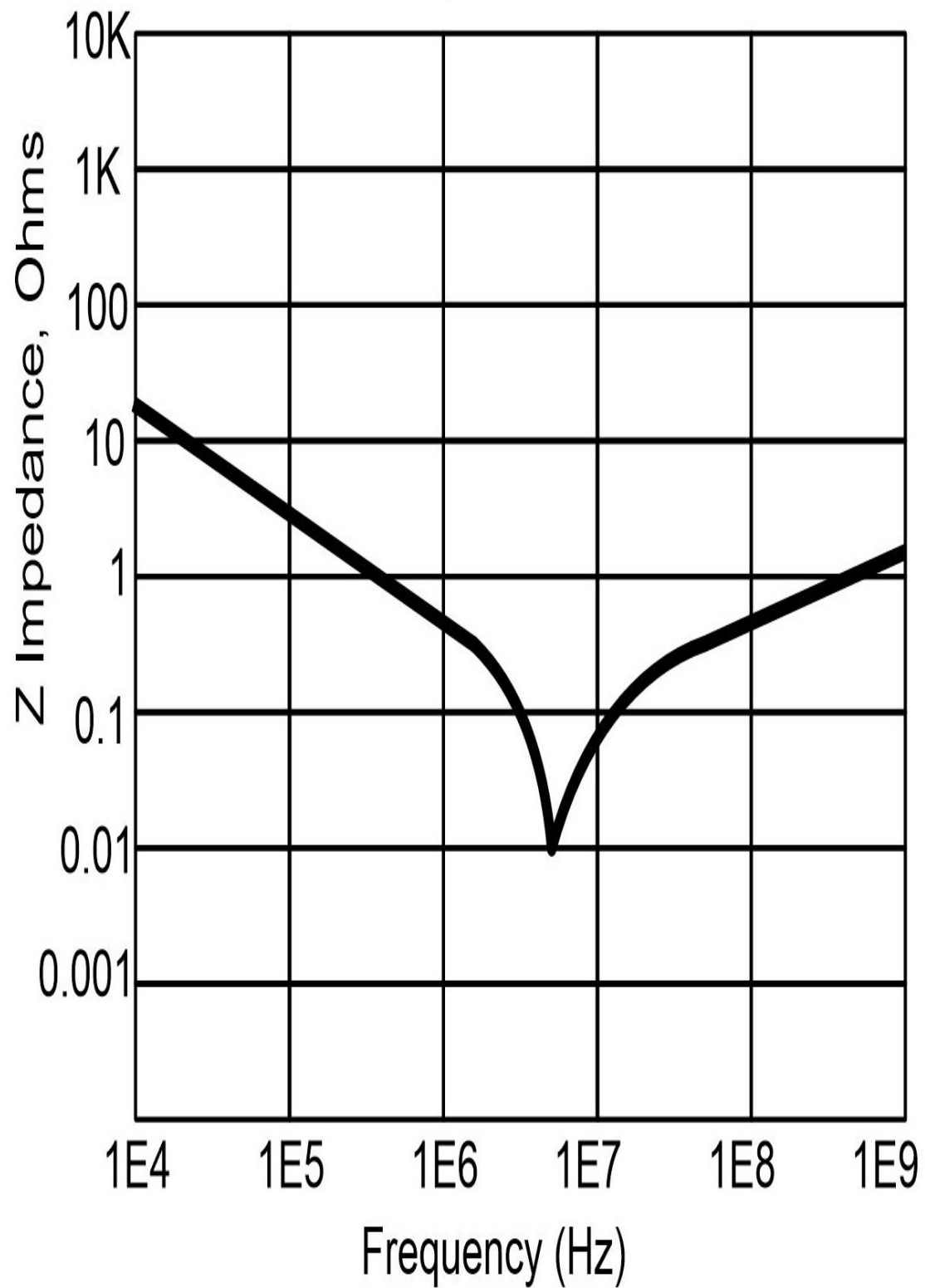


Figure 4-29. Self Resonance of Capacitors

Figure 4-? shows the self resonance characteristics of a typical surface mount capacitor. The low frequency response is dominated by the C_{main} capacitance, and the high frequency response is dominated by the ESL. In between, the device goes through self resonance frequency (SRF) and the impedance drops to a minimum which is equivalent to the ESR. The capacitor serves as a bypass filter beyond the SRF, but the effectiveness reduces with higher frequencies. The presence of ESL limits the high frequency capability of the device as a bypass filter for a power supply.

Self Resonance Frequency

C

SRF

10 μF	2.0 MHz
1 μF	5.5 MHz
0.1 μF	19 MHz
0.01 μF	57 MHz
1000 pF	180 MHz
100 pF	560 MHz

MLCC, SMT0603

Figure 4-30. Capacitor Self Resonance vs. Capacitor Size

Figure 4-? shows a typical set of values for the SRF as a function of capacitance value. Ideally, bypass filtering needs to avoid frequencies where ESL dominates the impedance. The reality is that many designs have wideband current demands that go beyond the useful range of most capacitors. There are conflicting rules of thumb and empirical studies of how to deal with this issue. (See suggested further reading material at the end of chapter.) In theory, using capacitors with different SRF points connected in parallel circumvents the SRF characteristics of singular devices is a common approach. This method is a common practice, but often blindly implemented without looking at the SRF of the capacitors. For this particular set of values, using capacitors two decades apart (such as 1uf, 0.01uf) would space the SRF in a suitable manner. Check the SRF of the specific devices used, because differences in manufacturer, dielectric type, or voltage rating, can have different SRF values. The SMT body size can also affect ESL and SRF with a SMT 0805 body having distinctly different characteristics than a SMT1206 body. Some vendors also offer “Low ESL” capacitors, as an option. These devices usually have wider contact regions and shorter bodies, or use multiple contact points to reduce inductance. Check the specific devices to be used for the details of their parameters.

Bypass Capacitors Value and Distribution

Bypass capacitors are utilized to do three separate things:

- A chip bypass is for the high frequency current demands of an IC.
- Grid bypass stabilizes local power grid connections, and reduces voltage variance due to both connection impedances and bandwidth limitations of the power source.
- Regulator Bypass filters the power source (regulator) output and stabilizes the feedback control loop within the regulator.

Semiconductor vendors should provide their suggested devices for at chip bypass to be used directly at their chips. If vendor information is not available, some guidelines are suggested here for at chip bypass:

- All power pins on all chips get dedicated bypass capacitors directly at the power pins.
- At chip bypass capacitors should have minimal trace routing and use shortest path connections. This is especially important with the wideband current demands of digital chips.
- PCB layout should avoid/minimize vias in connections of at chip bypass capacitors to minimize inductance.
- Nothing smaller than 0.01uF is suggested for at chip bypass. Although smaller capacitance values have higher resonance points, frequently the smaller devices don't exhibit low enough impedance, or are insignificant in the mixed collection of impedances of the power grid. (See Archambeault)
- If possible, keep the SRF of the capacitor 10X above the most common frequency used in the IC. High frequency chips will generally use the 0.01uF minimum as described above, some devices operating at lower frequencies may be able to use higher valued bypass capacitors at the chip.
- Use of X5R and X7R dielectric capacitors, in low impedance surface mount packages are preferred due to their limited capacitance change over voltage bias. Y5V devices exhibit extremely wide variance and can be unpredictable. The NP0 (or C0G) dielectric capacitors exhibit very little thermal variance and voltage bias variance. However, NP0 devices are limited in their capacitance values for a given volume device.
- Voltage rating of capacitors should be at least twice the maximum voltage seen in the circuit.

- Again, check with IC vendors for their preferred strategy for their chips.

Grid bypass capacitance is used for stabilizing the power planes locally.

Some guidelines:

- The local power plane is kept stable with distributed capacitance.
- The use of many capacitors spread out is better than a few capacitors.
- The capacitors used for grid bypass are typically larger value than the capacitor used for at chip bypass.
- For a multi layer PCB, put the grid bypass capacitors on the side of the PCB that is closest to Power/Ground planes. This minimizes inductance back to the Ground/Power planes.
- Distributed capacitance across the PCB should still use capacitors that exhibit good high frequency characteristics albeit resonate at a somewhat lower frequency than at chip devices.
- Using capacitors 100X larger in size than the at pin bypass devices keeps the resonance points well separated. For the situation where at pin bypass is 0.01uF, this implies grid bypass capacitors of 1uF.
- Grid bypass capacitors shall be distributed across the power planes such that any IC connected to the power plane has grid bypass in close proximity.
- Multiple grid bypass capacitors are suggested near the perimeter of any high gate count logic IC. (Microprocessors, Digital ASIC's, FPGA's, Memory IC's)
- A thin insulating layer in the PCB between Power and Ground planes may be used to provide some of the grid bypass capacitance. Use of embedded PCB capacitance does not remove the need for grid bypass capacitors, but can improve high frequency performance.

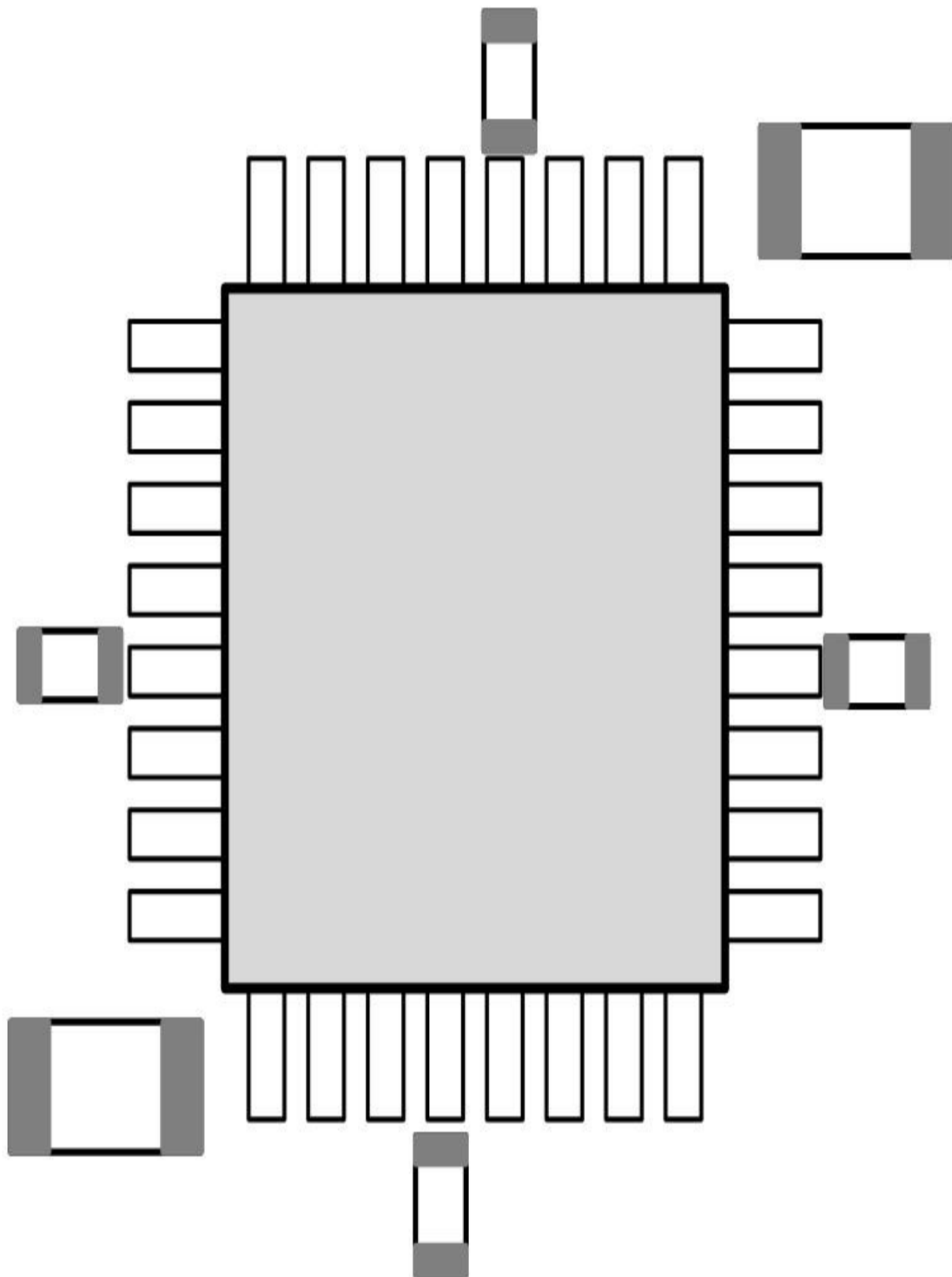


Figure 4-31. Bypass Capacitance, At Pin and Across Grid

Studies exist (Bogatin) that deal with grid bypass capacitance as controlled impedance vs. frequency problem. Circuit simulators can analyze the

interactive SRF of capacitors, but modeling the interconnection grid and everything connected to it is not quickly done. Typical design projects don't have the resources to make such a detailed study. These guidelines are an approach that can be implemented efficiently.

Regulator Bypass Guidelines:

- Manufacturers of any voltage regulator should be able to provide their preferred configuration for regulator bypass and decoupling.
- LDO regulators may need special attention to the ESR parameters of their bypass capacitors. Check the specifications and application notes carefully.
- Avoid electrolytic capacitors where possible due to their poor high frequency response and long term reliability issues.
- SMT MLCC devices are preferred over electrolytic capacitors for low ESR/ESL, reliability and size.
- Multiple SMT MLCC devices in parallel are a viable option for larger value capacitance.
- If electrolytic capacitors are unavoidable, placing parallel SMT-MLCC with the electrolytic capacitor will improve high frequency performance.

Essential Concepts – Conclusions and summary

The important points of designing a power system are:

- AC power requires a strategy to avoid electric shock and ensure safety.
- When the system fails, the device must remain safe. Understand how to avoid live contacts when in both normal operation and failure mode.

- Overcurrent protection needs to be part of a weakest link scenario. This is a must have for safety certification.
- Internal AC power needs grounding or double insulated safety methods.
- Offline switchers for AC-DC conversion are efficient and compact DC power sources. Look for devices that use internal self protection circuits.
- Regulate the power on the PCB where it is used. This minimizes problems due to connection impedance.
- Switching DC-DC converters can provide voltage reduction, voltage boosting, or both, in a compact and efficient manner.
- Linear regulators are inefficient, but useful where low noise on power, or low EMI in the environment, may be needed.
- Digital logic creates high frequency wideband surge currents on power and ground.
- Power and ground connections need dedicated PCB layers to create low impedance connections.
- Impedance of power and ground connections need to be minimized wherever possible.
- High frequency power stability requires bypass capacitors at all IC power pins.
- High frequency power stability requires distributed capacitance across the power grid.

After reading this chapter, the reader has the knowledge to...

Techniques for PCB layout of low impedance connections are covered in further detail in the PCB Design chapter.

Further Reading:

- “Protection Against Electric Shock” IEC 61140:2016
- “Audio/Video Information and Communication Technology Equipment – Part 1 – Safety Requirements” IEC 62368-1:2018:
- “Effective Power/Ground Plane Decoupling for PCB” Dr. Bruce Archambeault, IEEE EMC, October 2007
- “Switching Mode Power Supply Design”, 3rd edition, 2009, Pressman, Billings, Morey, McGraw-Hill ISBN 978-0-07-148272-5
- “High Speed Digital Design, A Handbook of Black Magic”, 1993, H Johnson, M Graham, Pearson, ISBN 978-81-317-1412-6
- “Decoupling Techniques” Analog Devices, MT-101 Tutorial,
- “An IC Amplifier User’s Guide to Decoupling, Grounding, and Making Things Go Right for a Change” Analog Devices, AN-202 Application Note, , Paul Brokaw.
- “Designing Second Stage Output Filters for Switching Power Supplies” Analog Devices, , Kevin Tompsett
- “Don’t Let Rules of Thumb Set Decoupling Capacitor Value”, EDN Magazine, Design Feature, September 1, 1995, Vincent Greb
- “Simple Grounding Rules Yield Huge Rewards”, Electronic Design Magazine, April 27, 2012, Jerry Twomey
- “Intel Pentium III Processor Power Distribution Guidelines” Application Note, April 1999
- “Signal and Power Integrity Simplified” 3rd edition, Eric Bogatin, 2018, Pearson Education, ISBN-13, 978-0-13-451341-6

Chapter 5. Battery Power for the System

Introduction

A NOTE FOR EARLY RELEASE READERS

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

This will be the 5th chapter of the final book.

If you have comments about how we might improve the content and/or examples in this book, or if you notice missing material within this chapter, please reach out to the editor at arufino@oreilly.com.

Modern electronics leans heavily towards mobile and portable devices, with the AC power cord relegated to battery charging. Battery power is a must have for modern devices, and even plugged in devices frequently have internal batteries to preserve critical systems when the power goes out. Many device fires, product recalls and explosions can be traced to improper battery system design. Also, many end users are frustrated by devices with insufficient battery capacity that can’t make it through a typical day of use. Clearly, there are a lot of marginal designs out there in need of improvement.

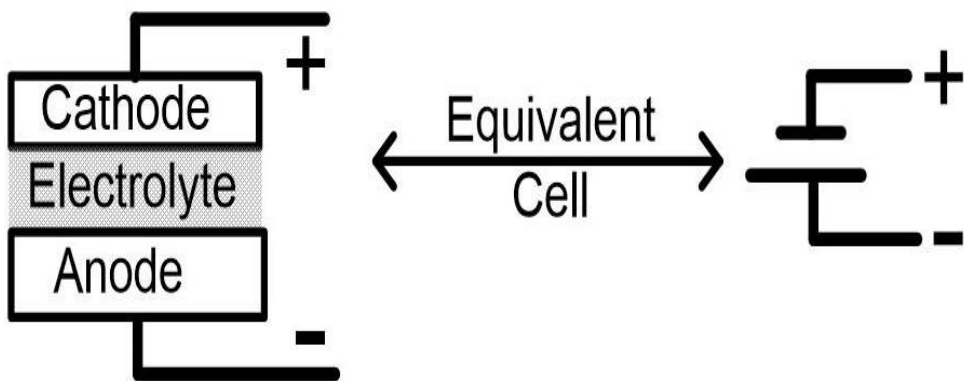
Most battery textbooks focus on the chemistry and physics of batteries with little published on incorporation into an electrical design. Designing a battery system to fit the needs of an electronic system is emphasized in this chapter with selection criteria and the needed support materials to make the proper design decisions. Chemistry discussions are minimal, largely treating

batteries as a black box, since readers are designing circuits and systems, not batteries.

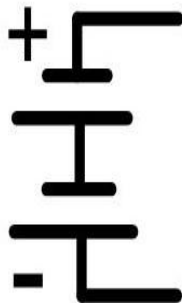
As the design progresses, make sure to obtain the performance specifications of the devices you intend to utilize. Battery technology is evolving and specifics need to be up to date.

Battery Basics - Definitions

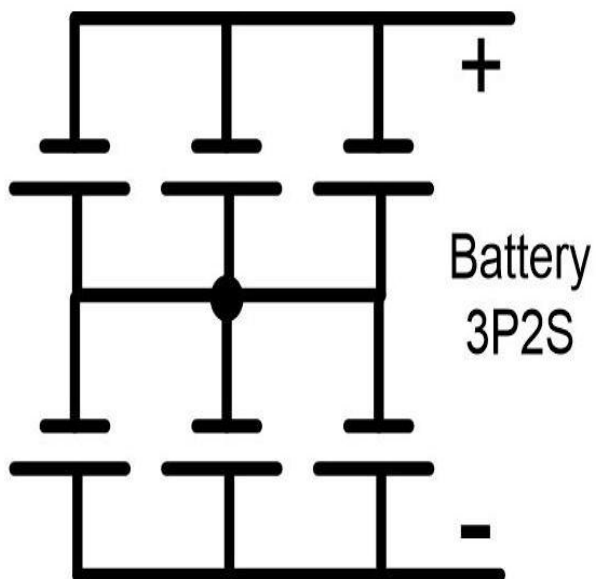
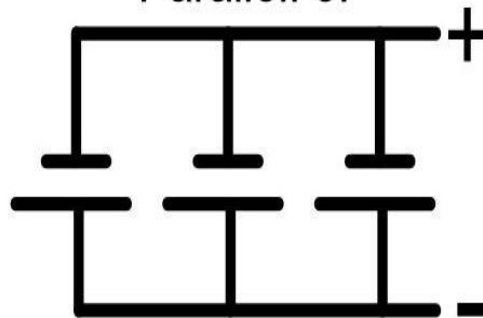
Let's begin with some definitions of the physical structure of batteries, and the terminology associated with device performance.



Series: 2S



Parallel: 3P



Generic Battery Symbol:



Figure 5-1. Physical Structure of Battery/Cell

The general structure of a single cell battery includes three distinct parts:

Anode

Electrical connection of the negative side of the battery. The anode emits electrons when the battery is in a circuit. The anode surface that is in contact with the electrolyte serves as an electrode.

Cathode

Electrical connection off the positive side of the battery. A cathode absorbs electrons when the battery is in a circuit. The cathode surface that is in contact with the electrolyte serves as an electrode.

Electrolyte

The electrolyte acts as a catalyst resulting in ions moving between the anode and cathode. The movement of ions, molecules that lack electrons, within the battery cause electrons to move in the external circuit connected outside the battery. The electrolyte is often soaked into a non conductive carrier material, known as a “**Separator**” to maintain physical distance between the Anode and Cathode.

Cell vs. Battery

A single “Cell” is the Anode-Electrolyte-Cathode trio as shown in Figure 5-?. Individual cells are combined in series or parallel to create a multi-cell battery. The terms battery and cell are often loosely interchanged. The common lead-acid 12V automotive battery is an example of a battery created with six series (6S) devices. The common AA battery, is actually a single cell.

Multiple cells are combined in series (2S) for higher voltage, or in parallel (3P) for higher current/capacity to create larger batteries. Combining cells in both parallel and series (3P2S) is commonly done.

The battery symbol has never been rigidly standardized; the number of plates in the symbol and whether the wider plate is positive or negative varies across the literature.

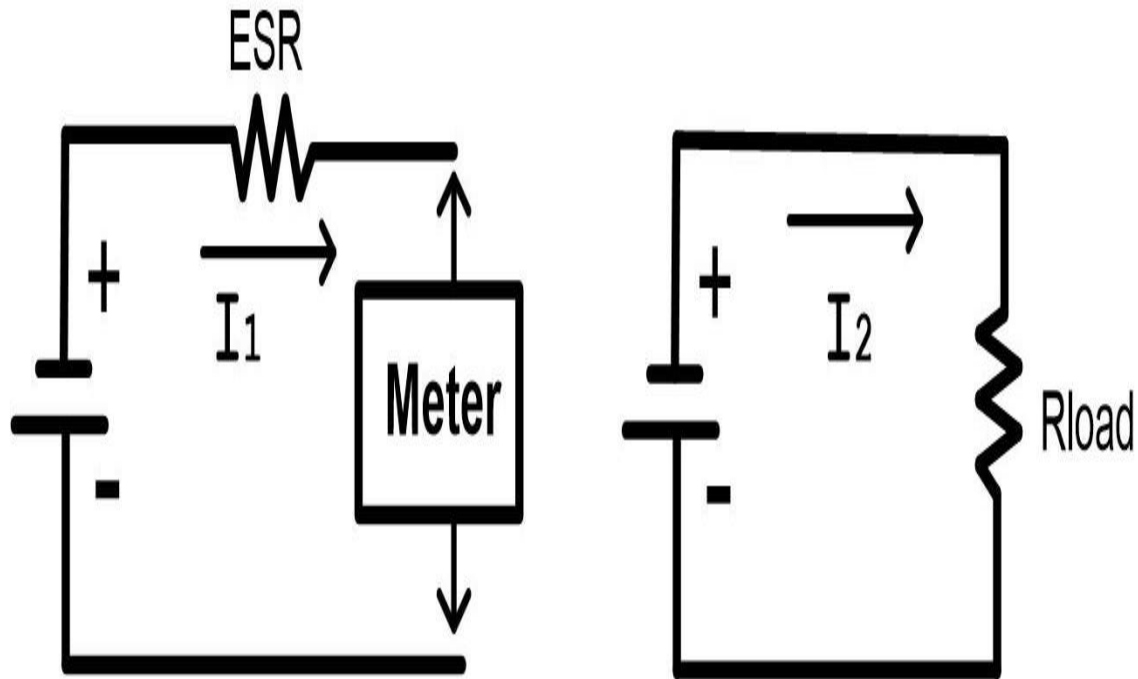


Figure 5-2. Electrical Criteria of a Battery

Open Circuit Voltage

Voltage observed on a battery output when the current is zero.

Internal Resistance

Equivalent Series Resistance (ESR) of the internal battery.

Capacity

Battery capacity is commonly measured in terms of current and time. The industry standard is to measure the capability to produce a fixed current and measure the duration. The units of (amperes x hours) are normally used. The amp-hour (Ah) is used for capacity comparison,

rather than how much electrical energy can be produced by the battery. Converting from Ah to Watt Hours requires multiplication by the average output voltage. Conversion to Joules ($1 \text{ J} = 1 \text{ Watt Sec}$), requires scaling by 3600 sec/hour.

C and C Rate

Capacity of a battery is commonly known as “C”. For an ideal device, the Ah rating is the same current needed to discharge the battery in one hour. The “C Rate” is a charge or discharge current that has been normalized to the capacity. If a battery has a 2 Ah capacity ($C = 2 \text{ Ah}$), putting a 2A load on it (load current = 1C) should discharge the battery in one hour. A battery with 55 Ah capacity requires a 55A load to discharge in one hour. Any battery should ideally discharge in 10 hours if the load is 0.1C.

Load

External element that discharges the battery. Battery loads are often expressed as equivalent current (a 250 mA battery load) or as a fraction of the C (a 0.01C load).

Charging Rate

This is magnitude of the current used to charge the battery. It is not the same thing as the C Rate. If a battery capacity is 10Ah and is charged using 1A, then the charging rate is 1A, or 0.1C. A battery with capacity of 600 Ah, and being charged using 60A, also has a charging rate of 0.1C.

Discharge Rate

Magnitude of current drawn from the battery while in use. Frequently this is normalized to the C value. If a 10Ah battery is discharging with 0.1A of current the discharge rate is 0.01C.

State of Charge

Defined as the present capacity of the battery as a percentage of the fully charged capacity of the battery.

Chemistry

The “battery chemistry” refers to the materials used to make the cathode, anode and electrolyte of the battery. Lead-Acid, Lithium-Ion, Carbon-Zinc, Nickel-Cadmium (Ni-Cd), Nickel-Metal Hydride are common names in use. The naming convention is not rigidly followed however; the “alkaline” battery is based on zinc and manganese dioxide.

Primary Battery

A battery is considered “primary” if it is a single use battery. In this text, the terms “single use” or “not rechargeable” are used.

Secondary Battery

A battery is considered “secondary” if it can be recharged. In this text, the term “rechargeable” is used.

Cylinder Battery

Battery built with an external shape of a cylinder. The name doesn’t indicate anything about the internal structure.

Prismatic Battery

Battery built with an external shape that has flat surfaces, often rectangular or square.

Battery Refresh

Used here, the expression “battery refresh” can be either the replacement of single use cells, or charging the batteries.

Charge Time

Elapsed time to recharge a battery.

Use Time

Elapsed operating time, between battery refresh events.

Shelf Life

How long an unused, properly stored, battery remains in good functional condition.

Service Life

This text uses the terms Use Time and Shelf Life. Depending on the information source, the term Service Life has multiple meanings.

Energy Density

Energy density serves as a comparison metric for how much energy is stored in a given volume or weight. Typically measured in (Watt)(Hours)/(Kilogram) or (Watt)(Hours)/(Liter). Depending if a system is constrained by weight or volume issues, either can be considered.

Power Density

Power density serves as a comparison metric for how much output power can be created within a given volume or weight. Typically measured in (Watts)/(Kilogram) or (Watts)/(Liter). Devices with high power density can briefly provide large amounts of energy to serve a particular application. As an example, an internal combustion engine uses a battery, optimized for high power density, to provide high amounts of “cranking current” to start the engine.

Decision Guideline – Rechargeable or Single Use Batteries

For some products, deciding whether to use single use batteries or the capability to recharge is self evident. Some situations are not straightforward. Some things to be considered while making that decision:

Product cost goes up when including a recharge system. Internal electronics are more costly with the charging circuits, adding an AC/DC source for the charger, the higher cost of rechargeable batteries and possible problems with the air shipping of restricted (Li-Ion) materials. Using snap in holders for single use batteries and adding “Batteries Not Included” on the package is a low cost solution.

Devices that use low power and require infrequent battery refresh are friendly to single use batteries. Things like smoke detectors, remote controls, calculators, electronic keys, and wall clocks, are all obvious single use candidates.

Devices used infrequently should be configured to use single use batteries, due to better shelf life and slower self discharge characteristics. This includes emergency devices (flashlights and radios) non hospital defibrillators (AED), and similar.

Easily swapped battery packs with rechargeable batteries fit a scenario where lightweight portability is necessary but heavy power consumption is still needed. The construction industry has embraced power tools supported by multiple battery packs avoiding the mobility limitations of power cords.

Internally installed rechargeable batteries work well when power demands require frequent battery refresh, and the typical use scenario for the device allows charger access. Tablets, Laptops and cell phones are common examples.

Devices that activate during AC power outages are best served by rechargeable batteries. This allows an automatic reset and battery recharge when mains power has been restored. Both uninterruptible power supplies and emergency exit lights fall into this category.

After deciding on single use or rechargeable, the next step is determining the power budget.

Defining Power Requirements

Electronic systems can have power consumption extremes depending on the application. Supercomputers run on megawatts and modern electronic watches function on 2-6 μW . Determining exactly where a design sits in this broad energy spectrum is an important part of defining the battery strategy.

Several things need to be determined:

- Use time
- Peak currents
- Long Term Average of Current over all use scenarios (LTAC)
- Functional voltage range of the system, min-max voltages needed

Use time is straightforward. Some examples:

- Smoke detector > 12 months
- Cell phone > 20 hours
- ZigBee devices mandate a use time > 2years for certification

One variation on use time is the “number of clicks” concept. For things like a wireless mouse, or a remote control handheld, specifying the number of clicks between battery refresh events is another approach. Devices need to consume negligible current between click events for this to be valid. Amp-hours per click is determined and multiplied by defined minimum clicks to determine needed battery capacity.

Defining peak current and LTAC require examining the transient current of the device. This can be done in the lab with a test fixture, or an accounting of all active circuits can be manually estimated.

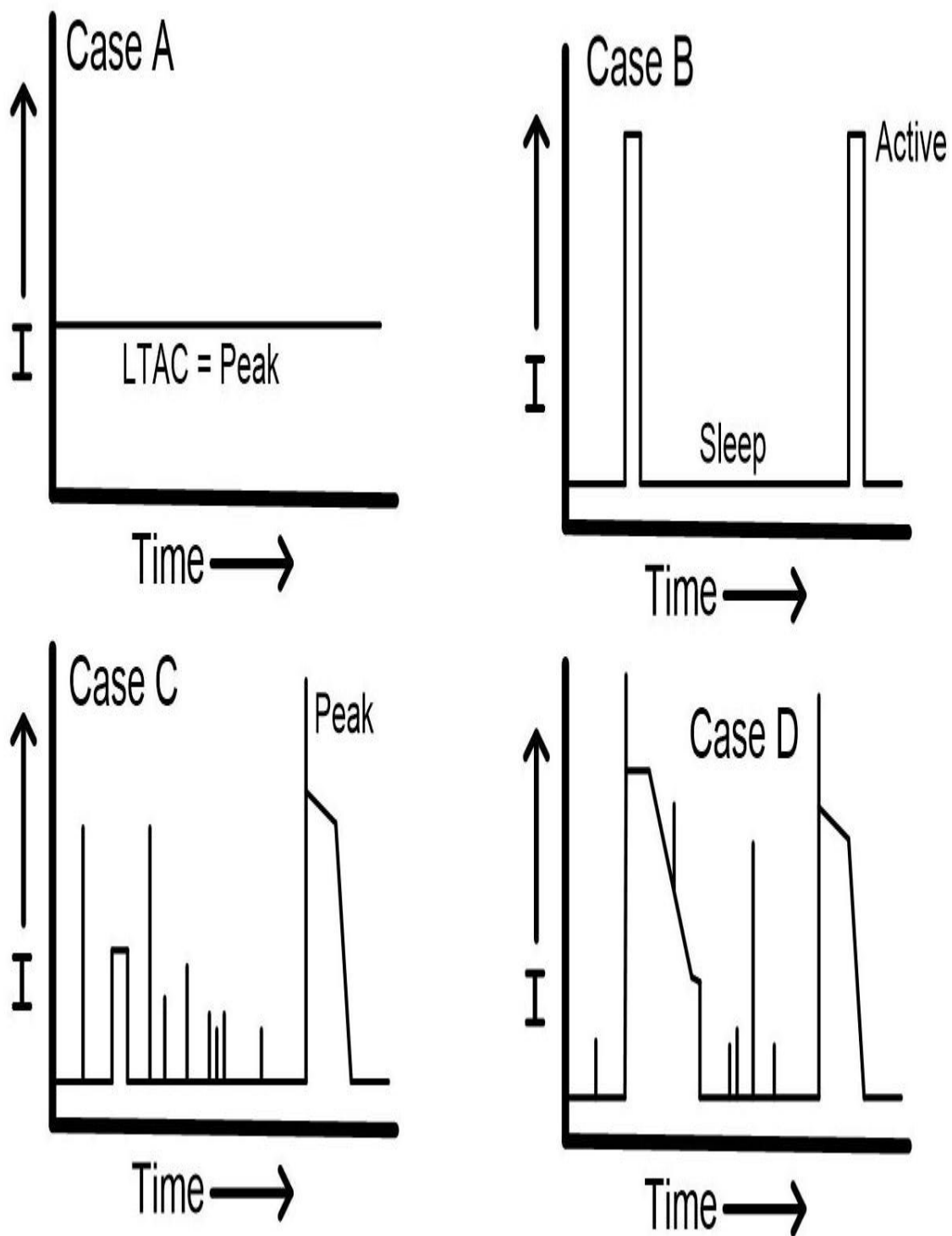


Figure 5-3. Power Budget Current Profiles

Depending on the system design, many different profiles of current consumption are possible. In this figure, Case A is the simplest scenario,

where a steady state current is the same as the LTAC. For Case A, the LTAC and peak current are the same. Case B is common for many single purpose systems; where brief bursts of high current activity are separated by long, low current, sleep periods. The LTAC will be highly dependent on the time periods of active vs. sleep and their respective currents. The peak current occurs during the active period.

Most multi-functional systems will have current use profiles that look like Case C or Case D. In both situations, various things turn on-off for intermittent periods. Prior to building hardware, a first pass estimate of LTAC can be based on the high current devices and their expected amount of active time. Electromechanical devices (motors, servos, solenoids, etc) tend to be the elephant on the list. Peak current surges are commonly seen at the startup of a motor or servo device. Display lights, high performance FPGAs, and multi core microprocessors can also dominate a power budget.

Min-max operating voltages of the system depend on the power system design (see power system chapter). If battery power feeds directly to active circuits, you are limited by the circuit's required voltage. Many power systems use some form of voltage regulator up front allowing a wider functional voltage range.

With power requirements defined (use time, LTAC, peak current, functional voltage range) a battery can be fitted to the electronics.

Battery Discharge vs. Functional Voltage Range

A typical battery discharge profile for a fixed current load is shown. The voltage output of the battery slowly descends, in a non linear manner, as the battery discharges.

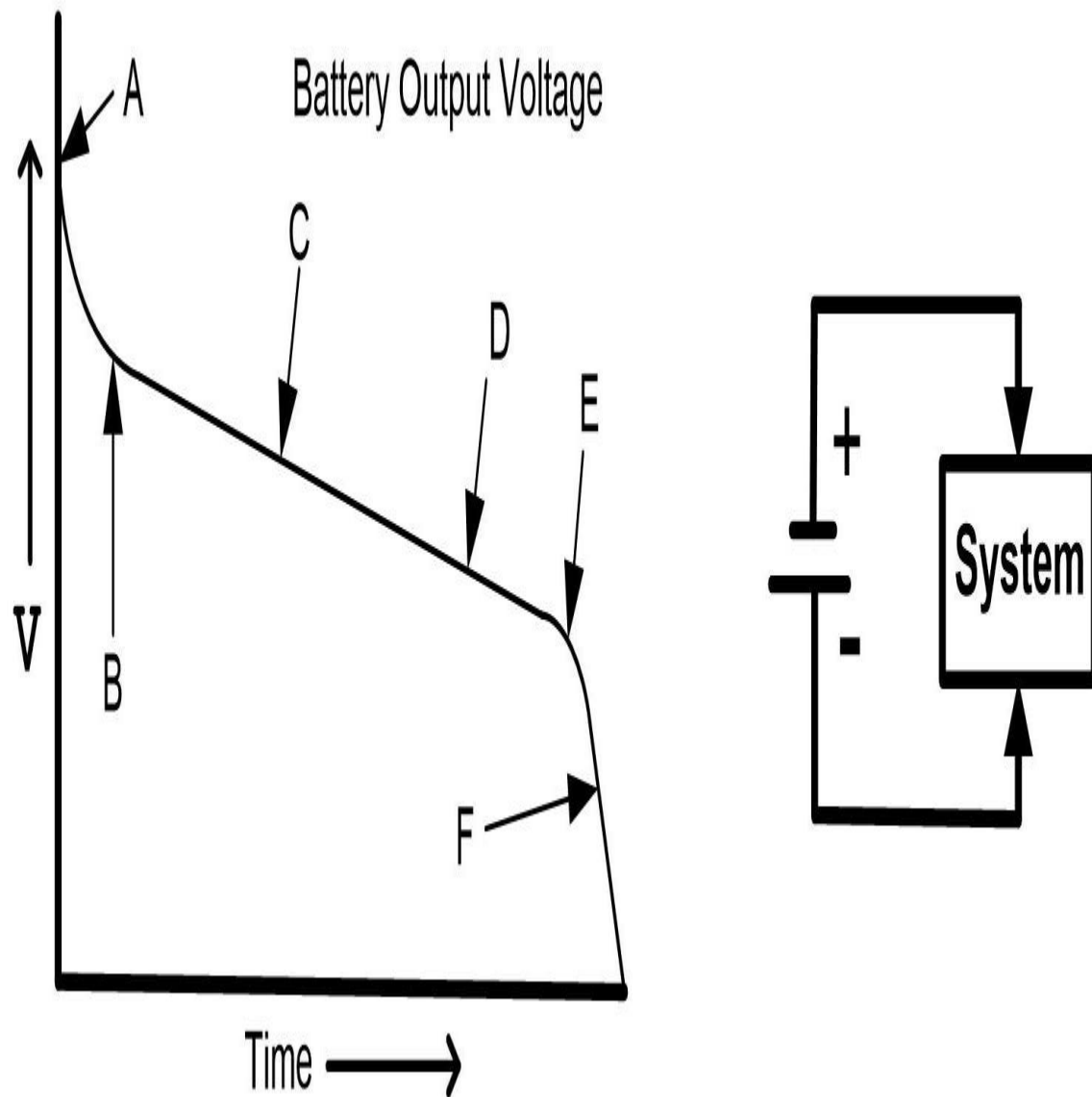


Figure 5-4. Battery Discharge Profile vs. Functional Voltage Range

The discharge profile starts from a fully charged device (A) where the battery has been “topped off” by the charger. Discharge quickly reduces the voltage to a region where a more linear discharge profile occurs (B, C, D) and discharge is considered complete at the knee in the curve (E). Discharge beyond (E) can sometimes damage the battery, depending on the chemistry. Single use batteries are often used down to (F), but many rechargeable battery management systems protectively cut off at (E) to avoid damage.

As the battery discharges, a progressively lower voltage is presented to the system. If the system can function down to (E, F) it is making full use of the battery's energy supply. If the system stops functioning back at (C, D) the full capacity of the battery is not being used. To utilize the full range of the battery, either a boost circuit can be used as part of the power system, or the battery pack can use batteries connected in series. Both methods are viable.

Battery Types by Chemistry

Different types of batteries have preferred features, with the division between battery types drawn along two lines, chemistry and package characteristics. Considerations when picking a battery include:

- Price and vendor availability
- Shelf life, (capacity deterioration, self discharge, physical leakage)
- Energy density (volume and weight)
- Recharge durability (# recharge cycles vs. capacity deterioration)
- Maintenance issues (water replenish, top off charging)
- Capacity (Ah)
- Peak current capability and Power density
- Safety restrictions for use and transport
- Capacity reduction due to temperature,
- Disposal restrictions and toxicity
- Charging time

Common Name	One Use	Shelf Life	Dis- chrg	Nom V/Cell	Min Off V	Features - Comment
Zinc Carbon (ZC)	S	4Y	10%/Y	1.5	0.8	AAA, AA, C, D, 9V variants available, low cost, low energy density, prone to leakage, performance very thermally dependent
Alkaline	S	7Y	3%/Y	1.5	0.8	Many variants available, 2X cost of ZC, moderate energy density, low internal resistance, performance thermally dependent
Li-Manganese Dioxide	S	12Y	1%/Y	3.0	2.0	Commonly used in Coin Cell batteries
Li-Iron Disulfide	S	12Y	1%/Y	1.5	0.9	AAA, AA variants available, 10X cost of ZC, good energy density, low internal resistance
Lead-Acid (SLA AGM)	R	5Y	5%/M	2.1	1.75	Many 6V, 12V variants, low cost, low internal resistance, good for bulk energy storage where volume/weight are not important
Nickel-Cadmium	R	3Y	20%/M	1.2	0.9	Flat discharge, fussy charge-discharge, low internal R, fast recharge, toxic disposal issues, replaced by Li-Ion in new designs
Nickel Metal Hydride	R	4Y	25%/M	1.2	1.0	Flat discharge, fussy charge-discharge, low internal R, suitable for high surge currents, lower cost than Li-ion, good energy density
Lithium-Ion	R	10Y	5%/M	3.6	3.0	Needs circuit protection & charge balancing, low internal R, suitable for high surge currents, very good energy density, safety issues
Lithium-Ion Polymer	R	?	7%/M	3.7	3.0	Similar performance to Li-ion, useful for non cylindrical batteries

S = Single Use 1%/Y = 1 percent per year
 R = Rechargeable 5%/M = 5 percent per month

Figure 5-5. Comparison of Battery Types Summary

Some additional comments on devices by chemistry:

Zinc-carbon

An older technology that largely survives due to low cost.

Alkaline (aka Alkaline-manganese)

Alkaline batteries are the high volume workhorse of single use devices. They are very popular for general use and a cost effective improvement over their zinc-carbon predecessor.

Lithium manganese dioxide (aka Lithium Metal)

Widely manufactured in coin cell package. Not readily available in other package variants.

Lithium iron disulfide (aka Lithium Metal)

High cost, not encouraged for new designs due to limited availability.

Lithium thionyl chloride (aka LTC)

Single use, very flat discharge curve, wide temperature range, good shelf life with low self discharge, high internal resistance so low current use only. As of 2020, LTC devices are starting to be made by more vendors. There are restrictions and safety warnings on LTC battery use. Check local regulations and restrictions before using in a design.

Lead-Acid – Flooded Cells

Commonly used in automobiles, trucks, golf carts and similar. Deep cycle variant uses internal plate structure optimal to deep discharge cycling. Cranking variant uses internal plate structure optimal to short period, high current surges. Not suggested for electronic systems due to slow charging time, maintenance issues, messy spill problems, and a need to be kept upright.

Lead-Acid – Sealed – AGM (aka Valve Regulated Lead Acid, Absorbent Glass Mat, VRLA-AGM or SLA-AGM)

The SLA-AGM battery holds the acid electrolyte in fiberglass between the cathode-anode plates. Sealed and no maintenance needed, SLA-AGM does not need to be held upright, making them popular in airplanes. SLA-AGM batteries come in many capacity sizes with 6V (3 cell) and 12V (6 cell) being the most popular voltages.

Lead-Acid – Sealed – Gel (aka Valve Regulated Lead Acid Gel, or SLA-Gel)

The Gel battery mixes the sulfuric acid electrolyte with a silica compound that results in stiffening of the mixture. No maintenance is required and the batteries can be used in any orientation. SLA-Gel and SLA-AGM have many of the same characteristics, with SLA-Gel having somewhat better thermal dissipation.

Nickel Cadmium (aka NiCd or NiCad)

NiCd batteries used to be a popular solution for power tools and other high surge current applications. NiCd devices have low internal resistance suitable for powering high current devices. More modern alternatives are now using Li-Ion and NiMH methods. Due to toxic Cadmium waste, some countries have banned NiCd batteries. NiCd is not suggested for new designs, due to toxic waste issues, and manufacturers are phasing them out.

Nickel-Metal Hydride (NiMH)

NiMH has low internal resistance and performs well with high surge current applications. NiMH suffer from large amounts of self discharge. With a poor shelf life they need to be kept on a charger when not in use.

Lithium Ion (Li-ion)

Li-ion batteries now have the bulk of the rechargeable battery market, with an established presence in everything from small electronic

handhelds to large automotive power cell arrays.

Lithium Ion Polymer (LiPo)

LiPo have similar electrical characteristics to Li-ion devices. LiPo devices use a laminated sheet method for the electrodes, separator and electrolyte, which allows devices to be made into many different shapes.

Global Behavior, Discharging

In order to better select your battery setup, it is useful to look a bit deeper at the characteristics of a battery discharging while being used. Simple Amp Hour ratings and nominal voltages can be very misleading as shall be shown.

First, determining the best approach to get quantitative answers on battery performance needs to be found. Does a math model for a battery exist, and can it be used in MatLab or Spice to get answers?

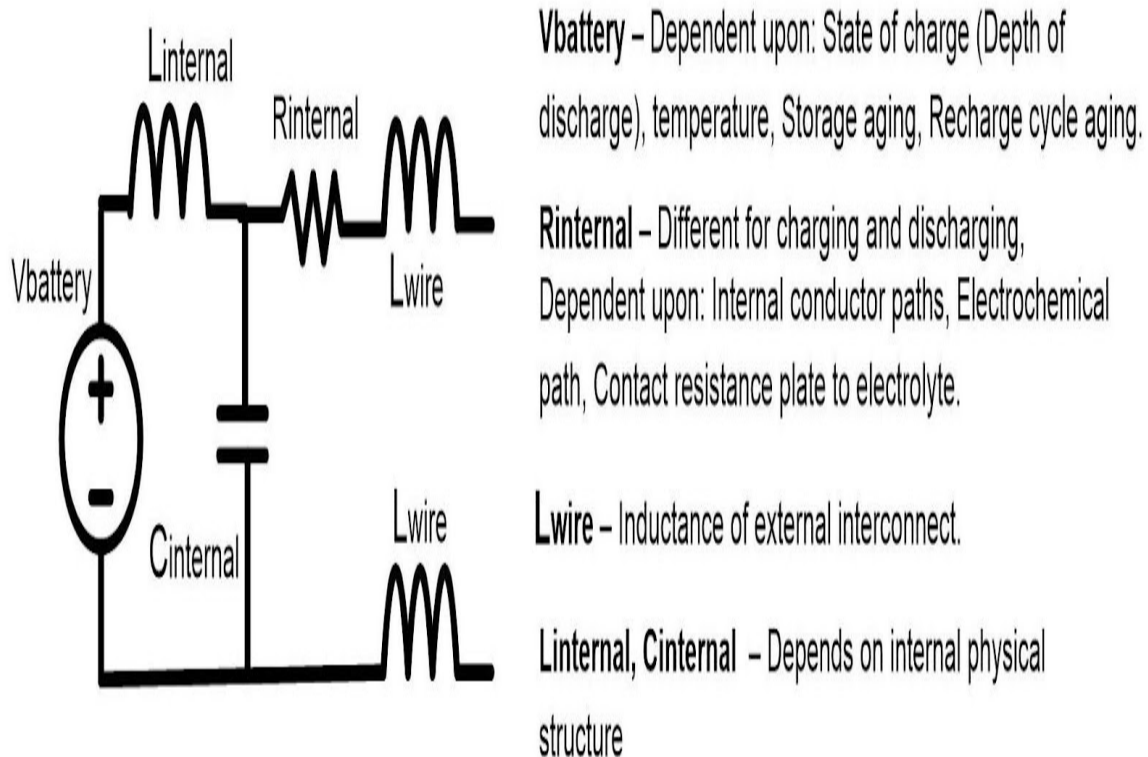


Figure 5-6. Battery Equivalent Circuit

A basic battery model would include some internal RLC impedance, a voltage source, and the interconnect inductance. For a DC only model, just R-internal and V-battery are needed. These are not fixed values however. R-internal changes for charging and discharging operations, and with aging of the battery. V-battery changes with state of charge, temperature of battery, aging due to storage, and aging due to recharge cycling. Simulation models can be developed, but require a lot of test data and curve fitting to create a meaningful model. Generally, it's not worth the effort.

Trying to simulate a battery doesn't expedite the design of a suitable battery pack. The battery discharge profiles can provide a much easier way to get a solution. Taking a typical AA alkaline battery as an example (Figure 5-?).

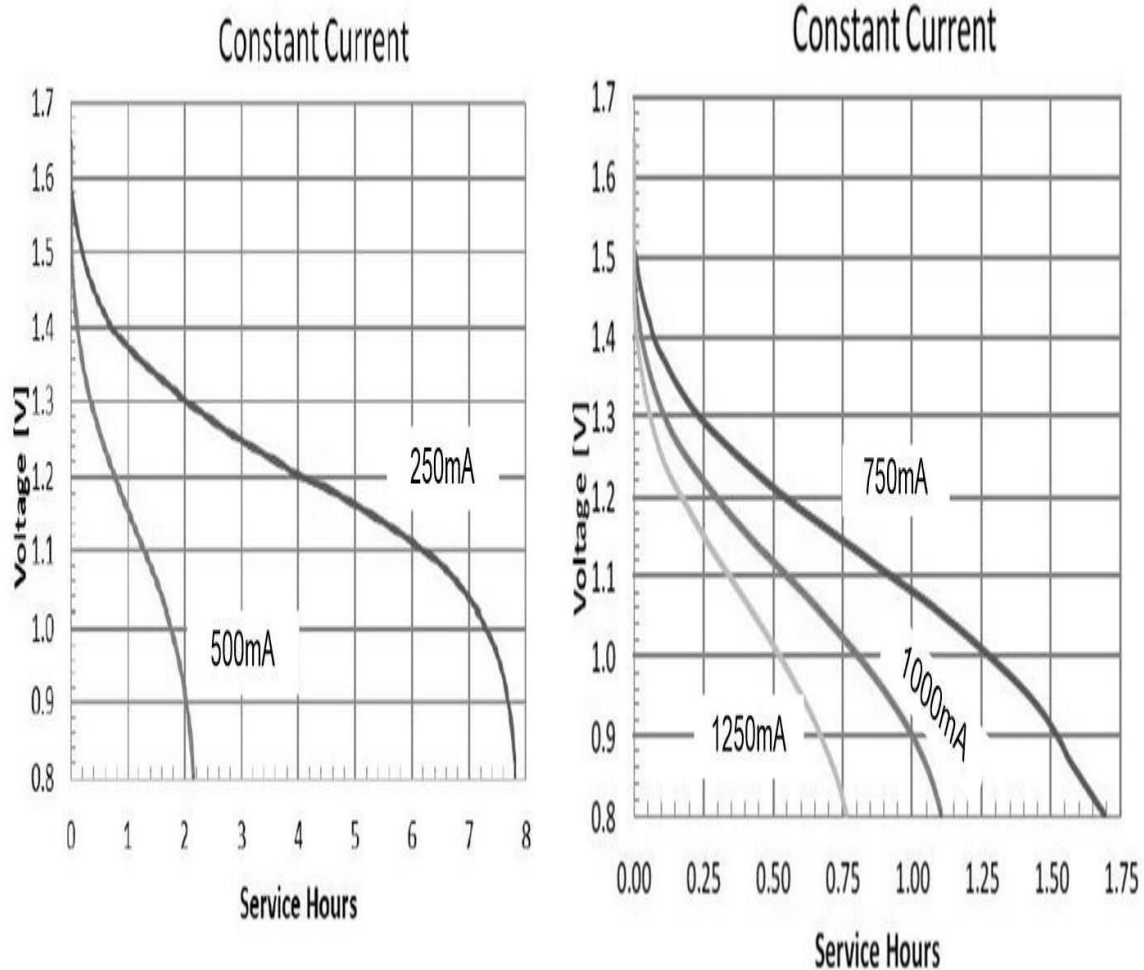


Figure 5-7. Battery Discharge Current vs. Use Time (AA Alkaline)

The above curve set shows the discharge profile of a common device for 5 different currents. Doing the math, the curves display Ah ratings between 0.9 Ah and 1.9Ah. Higher discharge currents generally produce a lower Ah rating. The stated battery capacity can be misleading. In this case, 2.7Ah was the stated capacity of the device. Vendors often state the Ah rating at a very low discharge current in order to get a better capacity number. Consequently, it is important to seek out further details of the capacity testing. Higher current means a lower Ah rating. Using a discharge curve for a current similar to the LTAC of the device should be done.

Most battery performance is tied to both discharge current and temperature.

Constant Current Performance

250 mA Discharge (-20°C / 0°C / 21°C)

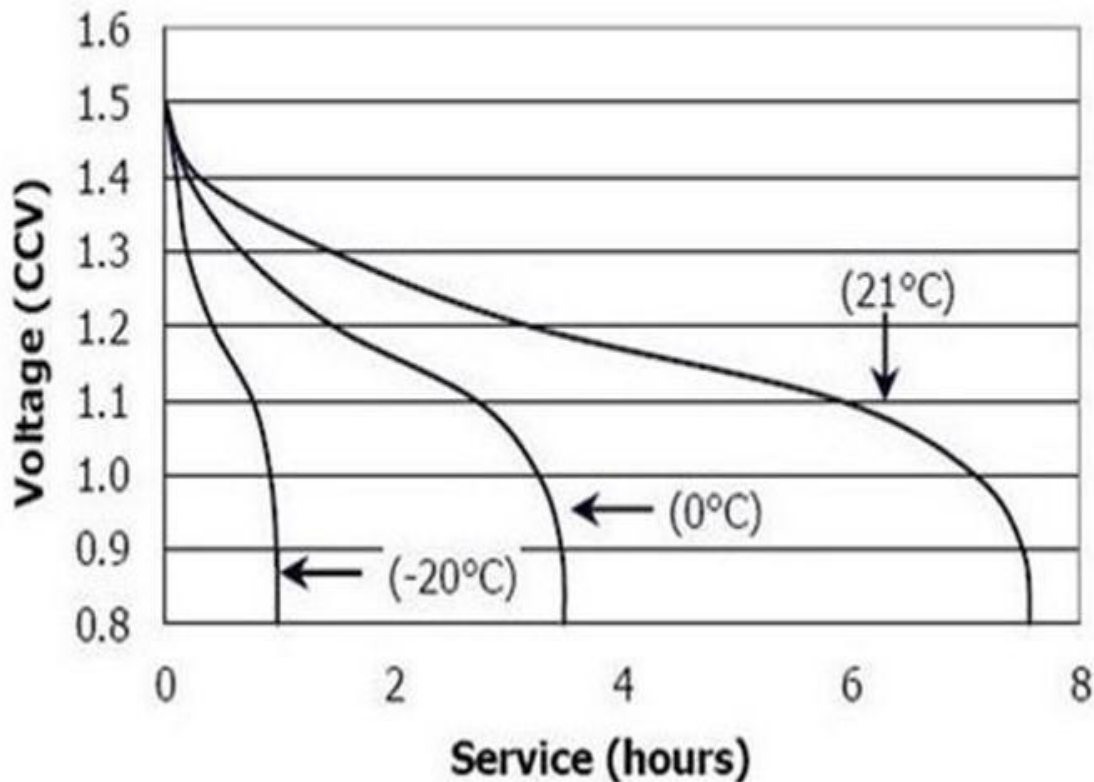


Figure 5-8. Battery Discharge Curves, Use Time vs. Temperature

As an electrochemical reaction, battery performance is very temperature dependent. This example shows a performance reduction from greater than 7 hours down to less than 1 hour over temperature extremes. The thermal performance of the battery and the temperature of the application environment are important. Many vendors test with warm batteries (25-30 degrees C) to produce better capacity numbers.

It is important to get the performance details and check the LTAC vs. Use Time over temperature. Using higher capacity devices, or, batteries in parallel, are the most common strategies to improve use time. Battery capacity is not a rigid number and margin needs to be included. Peak current also needs to be checked to avoid battery damage.

Combining batteries in parallel, or using the proper capacity device, get's an appropriate use time. The number of batteries in the series depends on the chemistry and the minimum voltage needs of the system.

Expected Voltage Range			
Type:	Discharge Min. Voltage	Nominal Voltage	Max Voltage
Alkaline E.	N(0.8) C.	N(1.25) D.	N(1.5) A.
Lead-Acid	N(1.75) C.	N(2.0) D.	N(2.4) B.
Li-Ion	N(3.0) C.	N(3.6) D.	N(4.2) B.
NiMH	N(1.0) C.	N(1.2) D.	N(1.6) B.

A - Voltage determined by use of new battery.

B - Voltage created by presence of battery charger.

C - Voltage limit defined by battery chemistry.

D – Dependent on nominal load and temperature, check specific battery documentation.

E – Common alkaline devices, AAA, AA, C, D, except 9V

N – Number of battery cells in series connection 1, 2, 3...

Figure 5-9. Expected Voltage Range

The table illustrates the expected useful voltage range from a single cell battery, broken down by chemistry. The common alkaline battery gets run down as low as 0.8V per cell, well over the knee of the discharge curve. Since the devices are single use there is no concern with doing the battery structural damage.

The 9V transistor radio battery is a variant of the alkaline, and uses 6 cells within its structure. So $6(0.8) = 4.8\text{V}$ is the discharged minimum voltage should be designed for.

The common automotive 12V lead acid battery, contains 6 cells, so $6(1.75) = 10.5\text{V}$ should be expected when full discharged, and a maximum of $6(2.4) = 14.4$ volts would be expected when a charger is attached.

A comparison of the minimum voltage for Li-Ion (3V) and NiMH (1V) is useful. Three NiMH cells are needed to get the same voltage as one Li-Ion device. These two devices are often compared for use, and a good comparison needs to include similar capacity and appropriate cell count. Alternatives include designing voltage boosting into the power system.

In all cases, the electronics need to function at the Min/Max voltages to make full use of the battery. If the system electronics shut down before hitting $V_{\min}$, consider either series batteries to bring $V_{\min}$ up, or a boosting voltage converter circuit in the power system (see power system chapter) allowing functionality at lower voltages.

Designing a Battery Set – Single use multiple cells

There are many single use batteries available. Single source devices should be avoided whenever possible. For instance, there are cases of products being halted because the battery vendor stopped making a specialty. Devices summarized here are all widely available.

Common Name	Volt	Size (Dia. mm)	Size (Ht. mm)	Capacity
AAAA	1.5	8	42	670 mAH @ 10mA
AAA	1.5	10	44	1300 mAH @ 25mA
AA	1.5	14	50	2900 mAH @ 25mA
AA (Li)	1.5	14	50	3500 mAH @ 25mA
C	1.5	25	50	8,000 mAH @ 25mA
D	1.5	33	60	16,000 mAH @ 25mA
9V	9	17 x 26	48	600 mAH @ 25mA
9V (Li)	9	17 x 26	48	750 mAH @ 25mA

Figure 5-10. Table – Single Use Batteries (AAA thru C, D, 9V)

All of the above devices are Alkaline, unless stated otherwise. One important note from the above table is the comparison of AA in Lithium-Metal to Alkaline. Although the Lithium Metal has ~20% better capacity, vendor prices are ~5X higher for the Lithium over alkaline. If higher capacity is needed, 2 alkaline batteries in parallel are more cost effective. The 9V devices have a similar capacity vs. cost difference.

Capacity (mAH) Changes due to Load Current				
Load I:	@ 25mA	@ 100mA	@ 250mA	@ 500mA
AAA	1,300	950	600	300
AA	2,900	2,500	2,000	1,500
C	8,000	7,000	5,000	4,000
D	16,000	15,000	10,000	7,000

Figure 5-11. Table – Battery Capacity vs. Discharge Current, Single Use Alkaline Batteries (AAA thru D, 9V)

A reminder that, the battery capacity reduces at higher currents. Published capacity numbers are often done using low current and warm temperatures to produce higher numbers. This table shows how capacity varies with load current.

Standard batteries for small applications are mostly coin cell batteries. These are also known as watch batteries, hearing aid batteries, or button

cells. The IEC standard defines 65 different versions of coin cell batteries in their standard. These are the most commonly used coin cells (Figure 5-?).

Common Name	Volt	Size (Dia. mm)	Size (Ht. mm)	Capacity
392	1.5	7.8	3.6	44 mAH
357	1.5	11.6	5.4	150 mAH
CR1616	3.0	16	1.6	55 mAH
CR1632	3.0	16	3.2	130 mAH
CR2016	3.0	20	1.6	90 mAH
CR2025	3.0	20	2.5	163 mAH
CR2032	3.0	20	3.2	240 mAH
CR2450	3.0	24	5.0	620 mAH

IEC 60086-3:2013 Defines 65 coin style batteries. Vendor specific names do not all follow the same standard, verify compatibility by dimensions and voltage.

Figure 5-12. Table – Single Use Coin Cells

There is no standard for labeling and numbering of coin cells. Some vendors use labeling that includes diameter and thickness. A CR2032 battery for example, is 20mm diameter and 3.2mm thick. Compare physical dimensions and voltage to determine compatibility. Also, make sure multiple vendors source the device, before committing to using it.

The chemistry used in coin cells varies. Alkaline and various Lithium-metal devices exist, Silver Oxide and Zinc-Air devices are also offered. Coin cells that used mercury and other hazardous materials have been phased out.

Designing a battery set using single use batteries needs to take into account:

- Long term average current (LTAC) or Burst Use (Ah per click)
- Peak Current
- Use Time, or total clicks before battery refresh
- Min-Max voltage of system vs. the battery
- Temperature of environment

A low power wireless thermostat is a good example for single use batteries (Figure 5-?).

ZigBee Enabled Wireless Thermostat

Battery Life: 2 years

Sleep Time: 60 S

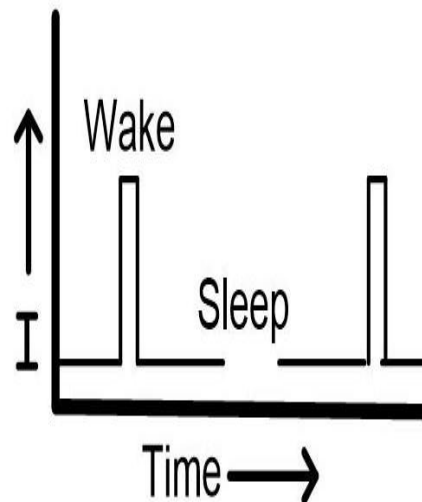
Sleep Current: 4.2 μ A

Wake Time: 250 mS

Wake Current: 9.3 mA

V_{max}: 3.3 V

V_{min}: 1.9 V



$$LTAC = \frac{(9.3E-3)(0.25) + (4.2E-6)(60)}{(60 + 0.25)} \quad LTAC = 43E-6 \text{ A}$$

Alkaline AA Battery: Capacity: 2900 mAh @25 mA @21 deg C

Use Time = $2900E-3 / 43E-6 = 67.4E3 \text{ H}$ Use Time is > 7 years

For 3.3 V_{max} to 1.9 V_{min}: Two AA cells in series will suffice.

Figure 5-13. Case Study – Single Use Batteries

The Zigbee standard requires devices have a battery life longer than two years. In this case, a thermostat wirelessly reports over the network every 60 seconds. The device goes live for 250 mS, using 9.3 mA, and then sleeps using 4.2 uA for 60 seconds. The LTAC is 43 uA, giving use time over 7 years for alkaline AA batteries, and over 3 years for AAA batteries. Since the operational $V_{\min} = 1.9\text{V}$, using 2 batteries in series is a suitable solution. The minimum voltage at full discharge is $2(0.8) = 1.6\text{V}$ so the batteries are not fully discharged but there's ample margin on the 2 year use time.

This device spends most of the time in a micropower (4.2 uA) sleep mode, so using a buck-boost converter to use the full energy of the battery is not suggested. The buck-boost converter would need to run all the time, and would greatly increase the current in sleep mode.

Designing a Rechargeable Custom Battery Pack

Picking the chemistry is the first step in the design of a custom rechargeable battery pack. Presently there are four contenders:

- Sealed Lead Acid (SLA)
- Ni-Cd
- NiMH
- Li-ion (including LiPo, Prismatic, and cylinder variants)

Of these, Ni-Cd is prohibited in some countries, and is being phased out by manufacturers due to toxic waste issues. By both mass and volume, SLA has the worst energy density, Li-ion has the best energy density and NiMH sits in between.

If low weight and small volume are top priority, Li-ion is the path to take. If weight and volume are not a concern SLA can be a low cost and reliable solution. High surge current capability is a strength of NiMH and has replaced Ni-Cd in many applications. Li-ion can be configured for high surge current scenarios but may need special safety shutoff circuits designed for it. If a special shape battery in a limited space is needed, a LiPo device is generally the best fit.

These are the most common offerings of SLA-AGM batteries (Figure 5-?).

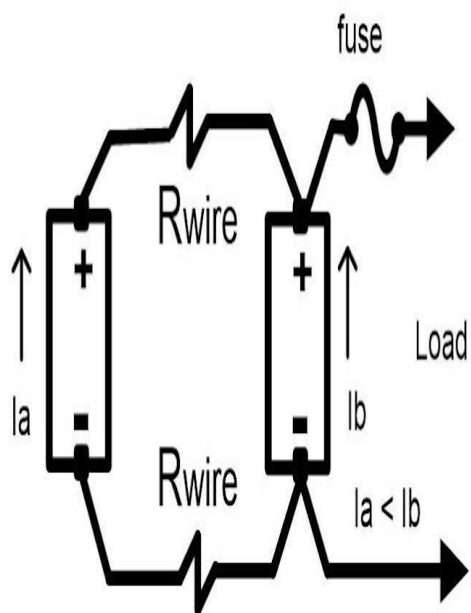
Lead-Acid (SLA, AGM) – Volt/Size/Capacity - Multi-Sourced						
Device	Volt	Height	Width	Length	Capacity	Weight
PS-612	6V	51 mm	24 mm	97 mm	1.3 AH	0.3 Kg
PS-640	6V	100 mm	47 mm	70 mm	4.3 AH	0.8 Kg
PS-670	6V	94 mm	34 mm	151 mm	6.7 AH	1.1 Kg
PS-6100	6V	94 mm	51 mm	151 mm	12 AH	1.9 Kg
PS-1212	12V	52 mm	43 mm	96 mm	1.4 AH	0.5 Kg
PS-1220	12V	60 mm	35 mm	178 mm	2.4 AH	0.9 Kg
PS-1230	12V	60 mm	67 mm	133 mm	3.4 AH	1.3 Kg
PS-1250	12V	101 mm	70 mm	90 mm	5.0 AH	1.6 Kg
PS-1270	12V	94 mm	65 mm	151 mm	7.0 AH	2.2 Kg
PS-12120	12V	93 mm	98 mm	151 mm	12 AH	3.6 Kg
PS-12180	12V	167 mm	76 mm	181 mm	17 AH	5.7 Kg
PS-12260	12V	125 mm	177 mm	167 mm	24 AH	7.7 Kg

Power Stream device numbers given, each vendor uses different numbering. Compatibility needs to be individually determined.

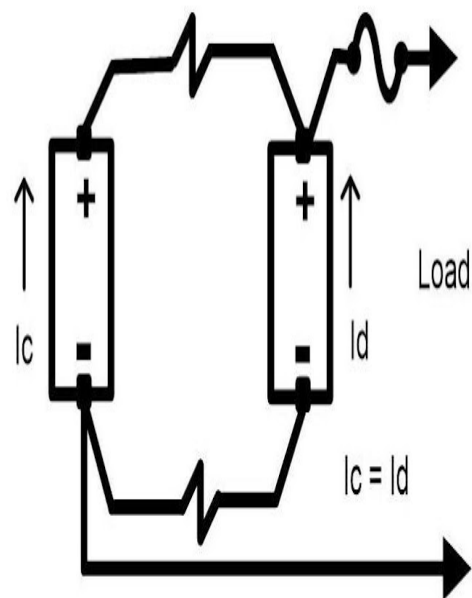
Figure 5-14. Table – Sealed Lead Acid (SLA)

SLA-AGM batteries are offered in many variants, most widely available are 3 cell (6V) and 6 cell (12V) devices. The above devices are widely available. Since there is no labeling standard for the devices, verify compatibility by dimensions, voltage, capacity, and connector location. A wide range of capacity is available and SLA devices combine easily in parallel and series to increase capacity and voltage.

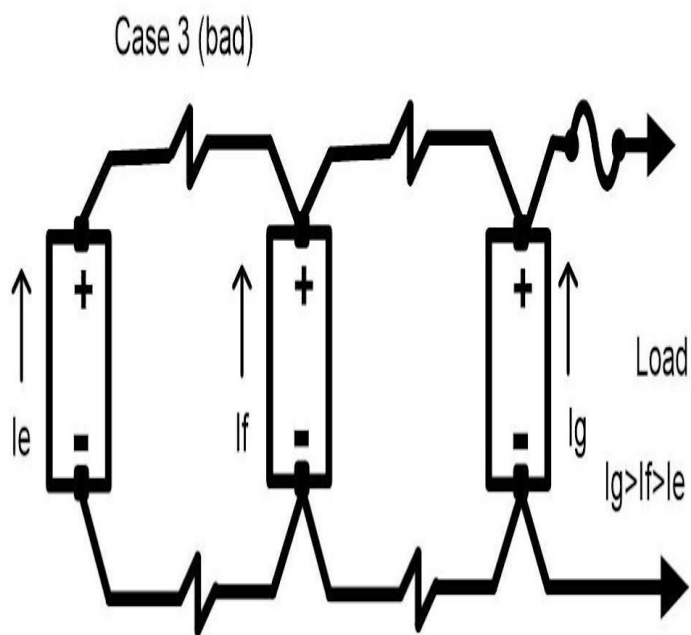
Keeping the battery currents balanced for both high current discharge and charging requires some attention to detail however. When combining batteries in parallel the IR drops in the wiring need to be kept balanced so the batteries provide equal currents to the load. The interconnect resistance and where it is placed can affect battery loading in high surge current scenarios.



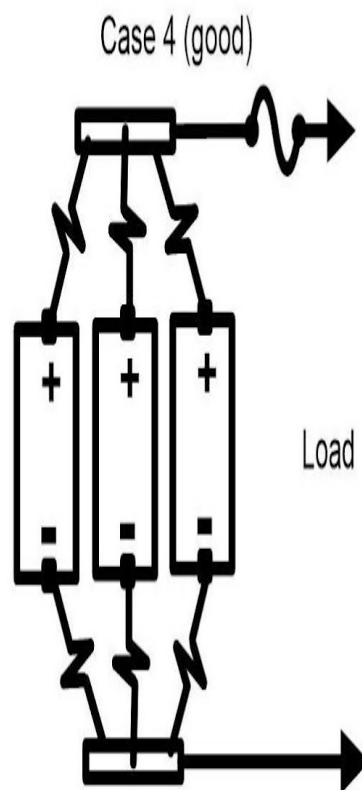
Case 1 (bad)



Case 2 (good)



Case 3 (bad)



Case 4 (good)

Figure 5-15. High Current Battery Connections

Case 1 illustrates a common way that two high current batteries are placed in parallel to increase capacity and surge current capability. When a high current surge is demanded by the load, the battery currents I_a , I_b will not equal because extra resistance is present from R_{wire} on both sides of the left battery. Because the batteries are connected in parallel, the charge state of the two devices will balance each other out when the load current is low. However, the battery on the right will provide a larger part of the surge current and will tend to degrade in performance quicker due to it.

In all cases, R_{wire} should be kept identical. In high current situations milliohms can be significant. Case 2 fixes the surge current imbalance by making sure each battery has a single R_{wire} in series.

Case 3 is an extension of the problem to three batteries, where surge currents will be different from all 3 devices. Case 4 rebalances the resistive paths to all batteries, by using tie together bus bars. In addition, it avoids stacking connections on battery connections which can be mechanically problematic.

In addition, when combining SLA devices keep the devices matched. Using the same age, same type/model, and same vendor gets the best balanced charge-discharge performance.

SLA devices will be individually purchased and manually configured into a connected array where needed. Battery packs using Li-Ion and NiMH cells are generally assembled by a specialized manufacturer.

NiMH for Custom Battery Packs				
Name	Usable Volt. Range	Diameter - Height		Capacity
AAA	1.4V – 1.0V	10mm	44.5mm	700 mAH @140 mA
AA	1.4V – 1.0V	14.5mm	50.5mm	1500 mAH @300 mA
A	1.4V – 1.0V	17mm	50.5mm	2450 mAH @490 mA
C	1.4V – 1.0V	25mm	50mm	3100 mAH @600 mA

Figure 5-16. Table NiMH for Custom Battery Packs

NiMH cells are available in a multitude of sizes. Some of the more common ones are listed above. Determine long term availability and specifics of performance before including in a design.

Lithium for Custom Battery Packs				
Name	Usable Volt. Range	Diameter - Height		Capacity
18650	3.7V – 3.0V	18mm	65mm	2000 – 3550 mAH, DOV
21700	3.7V – 3.0V	21mm	70mm	~4500 mAH, LA
20700	3.7V – 3.0V	20mm	70mm	~4200 mAH, LA
Prismatic	Custom, voltage range: N(3.7 – 3V), capacity as desired			
Pouch Cell	Custom, voltage range: N(3.7 – 3V), capacity as desired			

Labelling: <XX><YY><Z>

XX = Diameter, YY = Height

Z = 0 = Cylinder

DOV = Depends on Vendor

LA = Limited Availability

N = Integer, # series cells

Figure 5-17. Table – Li-Ion for custom battery packs

Availability of Li-Ion batteries is different than other devices. Custom shapes and sizes are available in prismatic variants (seen in cell phones, tablets, and other small devices). The predominant cylinder battery is the 18650 size. The 18650 is used in arrays from a single cell, to large (> 8000 cells) arrays in electric cars. The 18650 was widely manufactured for laptop battery packs, and as laptops slimmed down (requiring prismatic batteries) the devices became available for other applications at attractive prices. The larger 20700 and 21700 are starting to be built in quantity.

Due to safety issues with Li-Ion many manufacturers are reluctant to offer unprotected batteries on the open market. Instead, a common approach is to sell only to recognized manufacturers of battery packs, so that the devices are properly outfitted with safety circuits, charge balancing and support electronics.

Many battery pack vendors offer a multitude of pre-assembled battery array modules based on the 18650 with battery management electronics built in. Some of these come with safety regulatory certifications, and can be a cost/time saver.

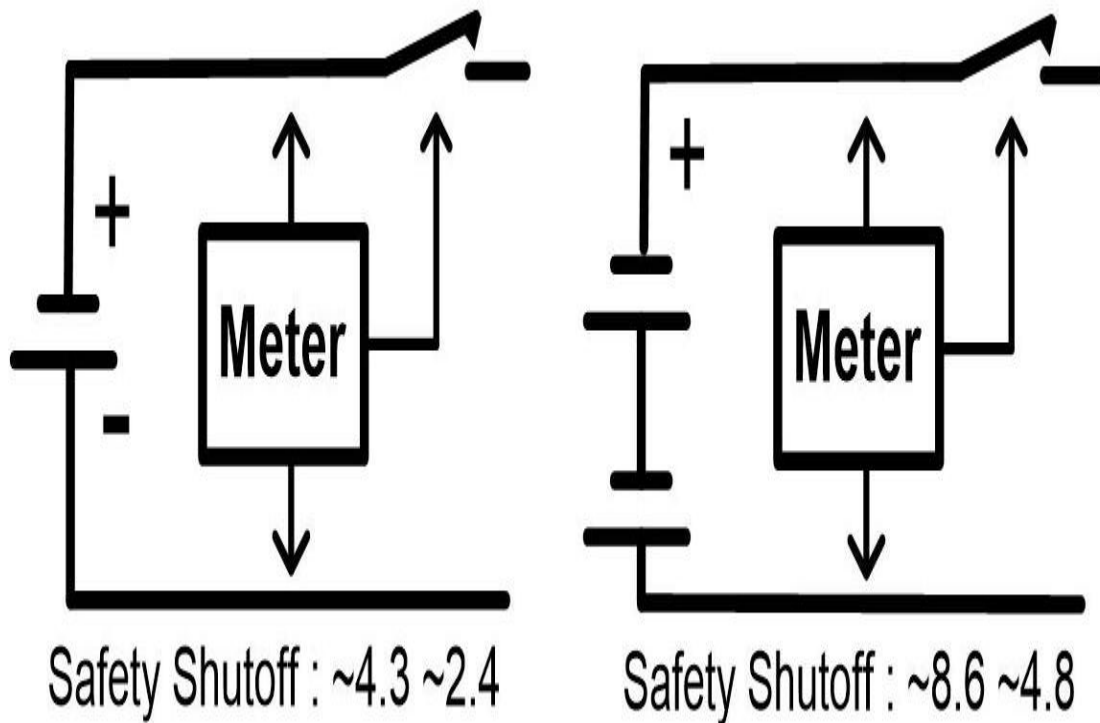


Figure 5-18. Li-Ion Safety Cutoff

Conceptually the safety cutoff device is simple: Disconnect the battery if the charger tries to exceed a maximum voltage ($\sim 4.3\text{V}/\text{cell}$) or the system tries to discharge the battery below a minimum voltage ($\sim 2.4\text{V}/\text{cell}$). For multiple batteries in series, the cutoff voltages scale up proportionally.

The safety cutoff is a last line of defense for the battery before damage or fire occurs. Charger systems, when properly configured, should stop charging about 4.2V , and power management should stop draining the battery around 3.0V per cell. The voltage limits mentioned here (4.3 , 4.2 , 3.0 , 2.4V) are typical, and most vendors will have similar values. If planning to multi-source batteries, a set of limits to accommodate all vendors can be worked out.

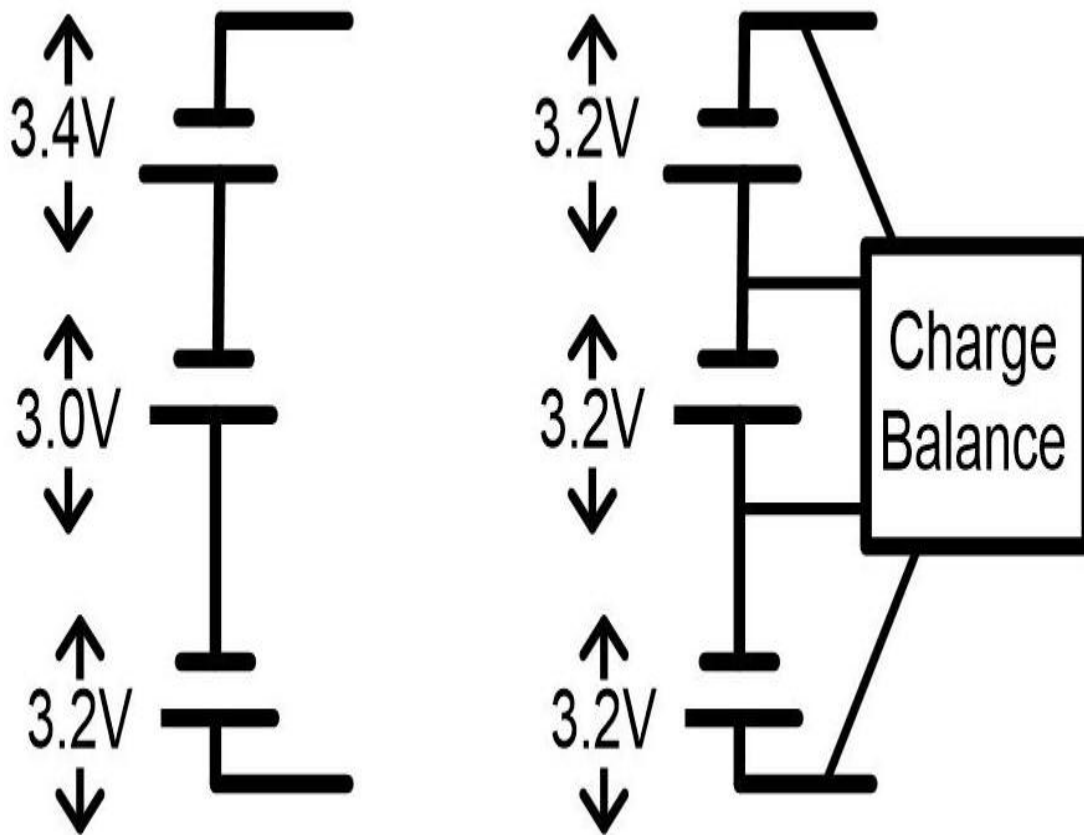


Figure 5-19. Charge Balancing

As emphasized throughout this book, nothing is ideal, including batteries. As illustrated on the left, when a battery with multiple series cells is charged, voltages are different across each cell. As the output voltage rises during charging, the charging of multiple series cells can be quite different. Charge balancing is needed to both maximize charge, and avoid individual cells from going into overvoltage.

On the right, a controlled bypass circuit connects around each cell. While charging, this keeps all cells at the same voltage, and all cell voltages rise together. Charge balancing circuits are readily available for cells in series from 2S to 32S.

Designing a battery set using rechargeable batteries needs to take into account:

- Long term average current (LTAC) or Burst use (Ah per click)
- Peak Current
- Use Time, or Burst use (total clicks) before battery refresh
- V_{min} & V_{max} needs of the system vs. V_{min} & V_{max} capability of the battery
- Temperature of environment and resulting capacity changes of battery
- Reduction in capacity due to aging
- Safety cutoff circuits
- Charge balancing circuits

As an example of specifying a custom rechargeable battery for a smart phone, this case study determines needed battery capacity by testing the various modes of the phone for current consumption (Figure 5-?).

Smart Phone Battery Use – Custom Prismatic LiPo

Use Time: 24 hours	Power Used While:
Vmax: 4.3 V	Suspend Mode = 70 mW = 21 mA
Vmin: 2.7 V	Lit Display = 300 mW = 88 mA
Vnom: 3.4 V	Web Browsing = 430 mW = 126 mA
Temp: 21 C	Active Calling = 1050 mW = 309 mA
	Texting = 300 mW = 88 mA
	Idle = 270 mW = 79 mA

Activity per 24h

Suspend Mode:	8h	8(21) mAh = 168 mAh
Idle:	10h	10(79) mAh = 790 mAh
Active Calling:	2h	2(309) mAh = 618 mAh
Web Browsing:	3h	3(126) mAh = 378 mAh
Texting:	1h	1(88) mAh = 88 mAh
Lit Display:	6h	6(88) mAh = 528 mAh

Total: 2,570 mAh

Rechargeable battery, specify a usable system with up to 20% loss of battery capacity as the battery ages:

$$2570 \text{ mAh} = (0.8) \text{ Capacity}$$

Capacity = 3200 mAh @400 mA
(current of: active call + lit display)

Figure 5-20. Case Study - Capacity Calculations for Rechargeable Battery

After determining the current used in each mode, the time for each mode is estimated for the use time. With a smart phone, several activities can be occurring simultaneously. The total Ah are summed and pro-rated for 20% loss of capacity due to aging. The results are typical to batteries used in smart phones presently on the market.

Charging Batteries

Battery charging methods are unique to chemistry. Each manufacturer will probably suggest a slightly different set of charging parameters to get the most performance out of their devices. In addition the number of cells in series, and parallel, scale the parameters of the charging process.

There are many custom IC's designed to control battery charging. Most are defined by the chemistry of the battery they are designed to charge. The parameters of charging are either digitally selected, or controlled using external components, to set voltages and currents needed for a specific battery pack. Multiple battery charging IC's are available for Li-Ion, Lead-Acid and NiMH devices.

Designers have done software defined charging by a host controller. However, high voltage controls are needed for the power transistors, voltage and current sensing are needed, and the dedicated charging IC is often a more cost and time effective solution.

Two examples of charging profiles are discussed here, one for Lead-Acid and one for Li-Ion devices. The information is for one cell. As devices are added in series, the voltage scales, as device are added in parallel the current scales.

Most batteries use constant current charging, where a fixed current is injected into the battery. The magnitude of the current is determined by the capacity and the chemistry of the battery.

Typical Lead-Acid Charging Profile

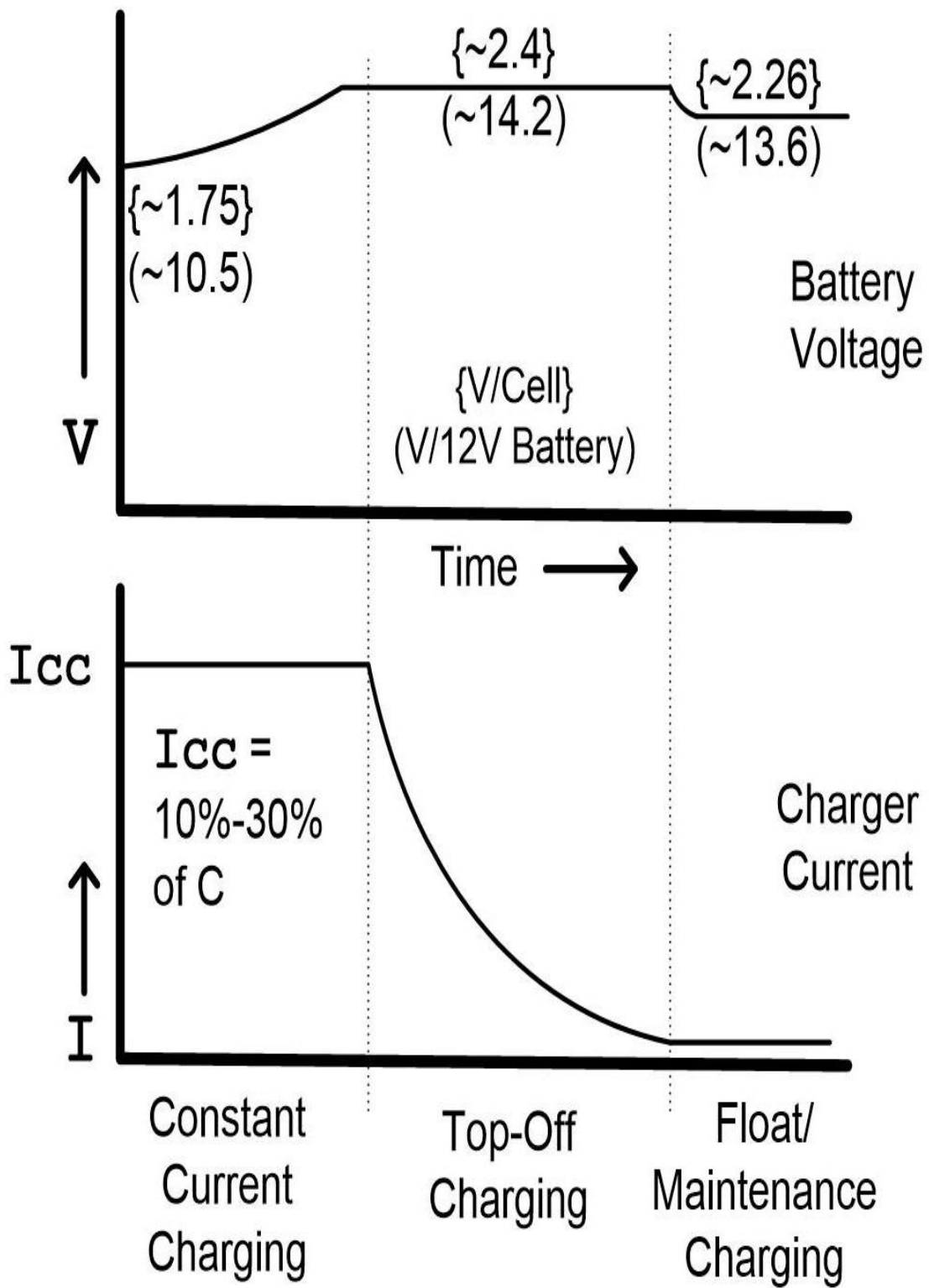


Figure 5-21. Lead Acid Charging Profile

A lead acid constant current charging strategy is shown. The constant current used, I_{cc} , is typically set at 10-30% of C for the battery. Higher currents would cause more energy to be lost as heat while charging. The current limitation implies that a lead acid battery takes 4-10 hours to fully recharge. Vendors will specify the details of their charging profile and the highest currents that can be used without doing damage.

A fully depleted cell is 1.75V or lower. When the cell rises to 2.4V, the current is tapered off in a segment referred to as “Top Off” charging. After top off, the charger holds the cell at ~2.26V in “Float” or maintenance mode until removed from the charger.

For a typical 6 cell automotive battery, depleted is 10.5V or lower, charging top off is 14.2V and float is 13.6V

A Li-ion battery has some differences (Figure 5-?).

Li-Ion 18650 Charging Profile

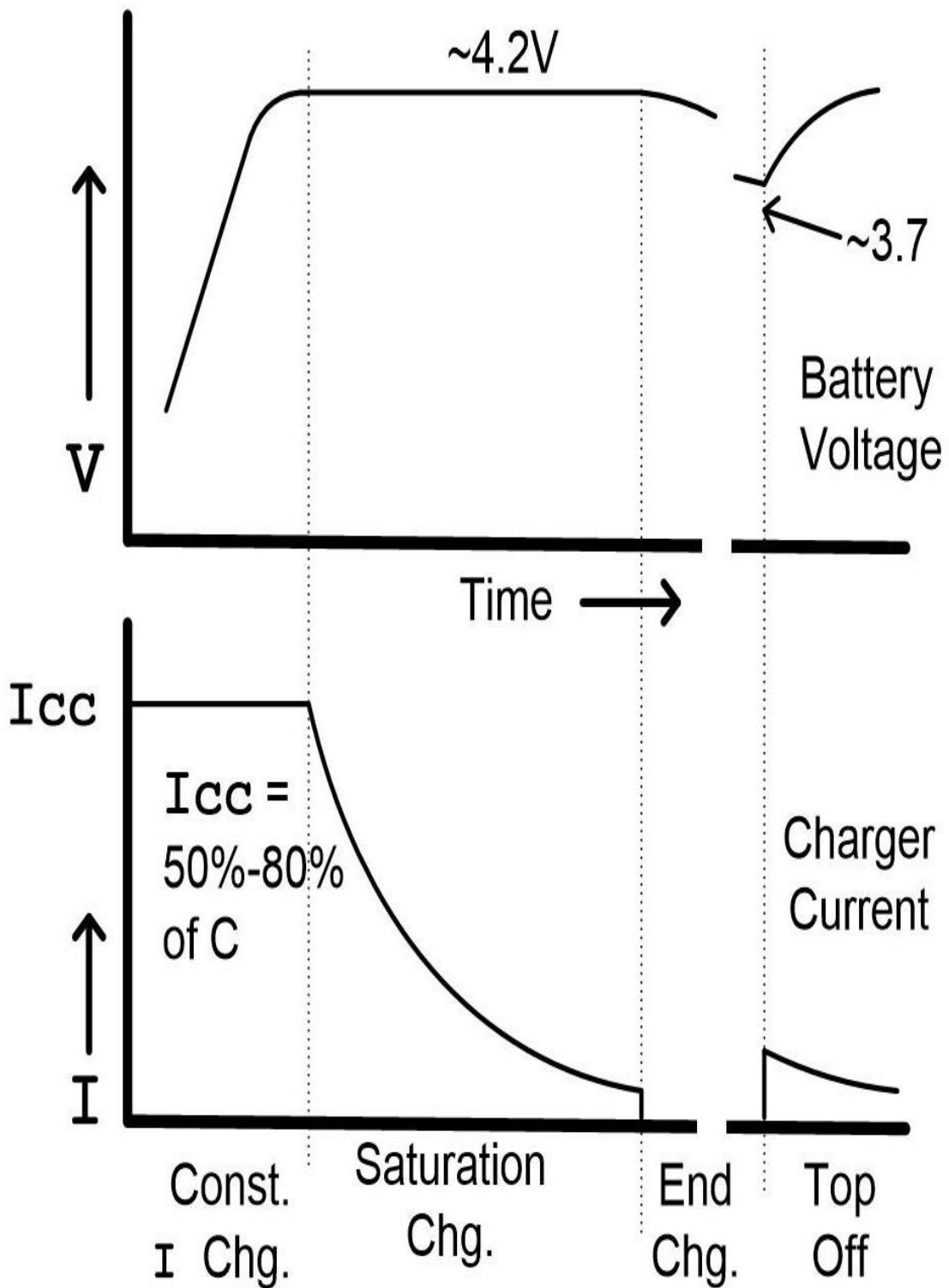


Figure 5-22. Li-Ion charging profile

Li-ion starts charging in a constant current mode. Most Li-ion devices use a charging current of 50-80% of C so a full recharge needs 1-3 hours. Again, the numbers are vendor specific.

A fully depleted Li-Ion cell is 3V or lower, and when the voltage of the cell rises to 4.2V the charging current is tapered down (saturation charging), and then turned off.

A maintenance voltage (float voltage) is not maintained on a Li-Ion battery. Rather the output is monitored until the device reduces to a defined point (typically 3.7V) where the saturation charging is re-enabled to top off the cell. Multiple cells in parallel scale up the C of the battery pack and multiple cells in series scales up the voltage.

Smart Batteries

An integrated battery pack can have features beyond than the battery cells. A battery pack with added features is often referred to as a “Smart Battery.” Consider a full feature example (Figure 5-?).

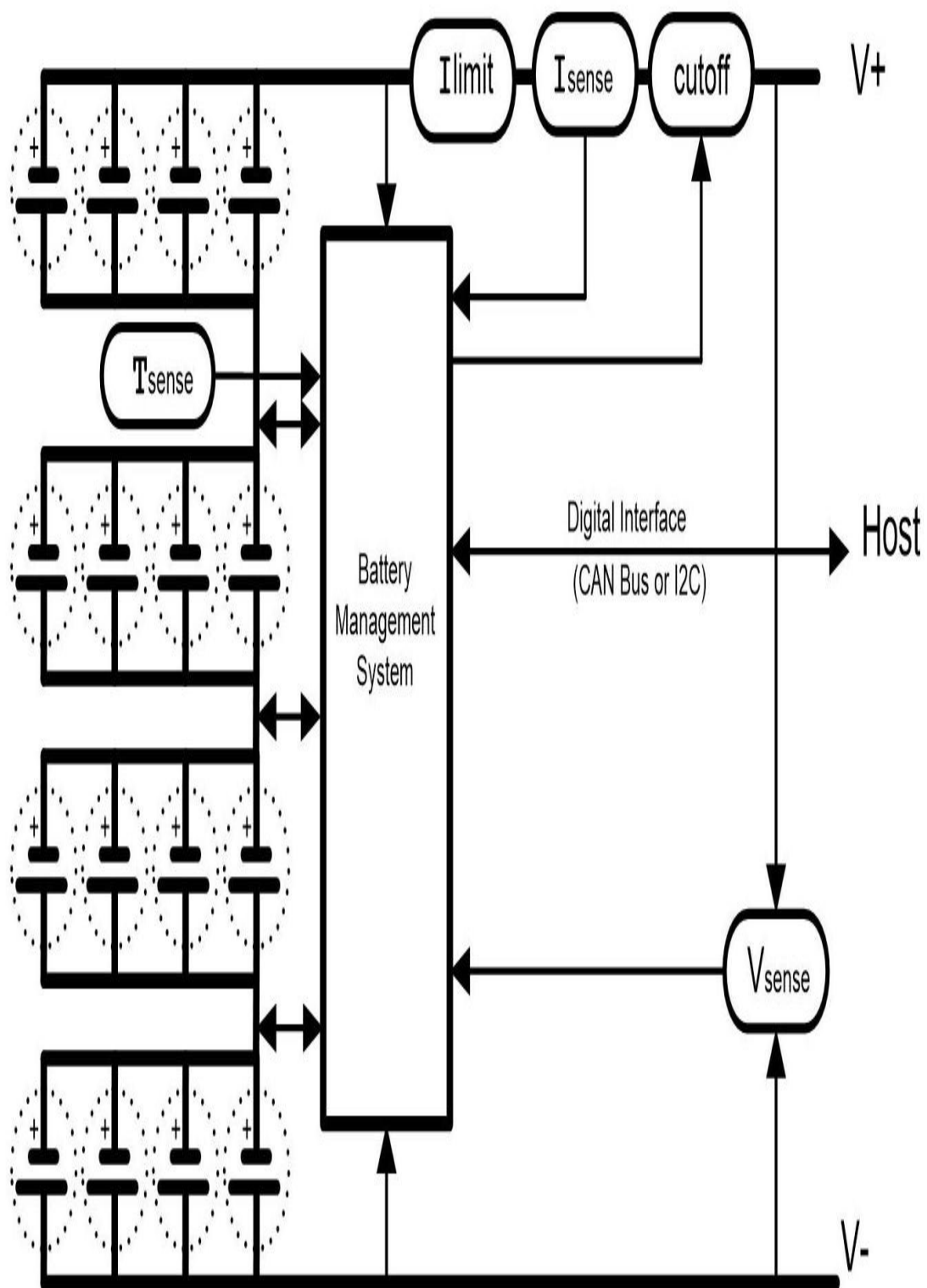


Figure 5-23. Full Feature Smart Battery System

The smart battery pack shown here has a 4P4S configuration. Cells are connected in parallel groups of 4 before the groups are connected in series to create the set. Connections from all 5 battery nodes are connected to a battery management system (BMS) that facilitates charge balancing and monitors/controls charging.

Thermal monitoring (T sense) is often included, especially within large arrays. Current limiting (I limit) can be done by using a thermally resettable fuse (aka PPTC) or by controlling the cutoff switch through the BMS.

Current sensing (I sense) allows host feedback on both charging and discharging rates. A safety cutoff (cutoff) prevents overvoltage (4.3V/cell) and under voltage (2.4- 2.7V/cell) from happening. Voltage sensing is also there for the whole battery pack.

The voltage sensing and current sensing can be used to provide state of charge monitoring for the system. Time, voltage and current information can be monitored by the host process to create a coulomb counting system, or a “Fuel Gauge” status can be estimated by the overall voltage of the battery. Both are commonly used.

Digital serial interface control is available back to the host process commonly thru an I2C/SMB interface. Future devices should migrate to a CAN bus for better data integrity and error checking. Of course, the smart battery still needs an external battery charger to provide the needed charging profile to support a 4P-4S battery pack.

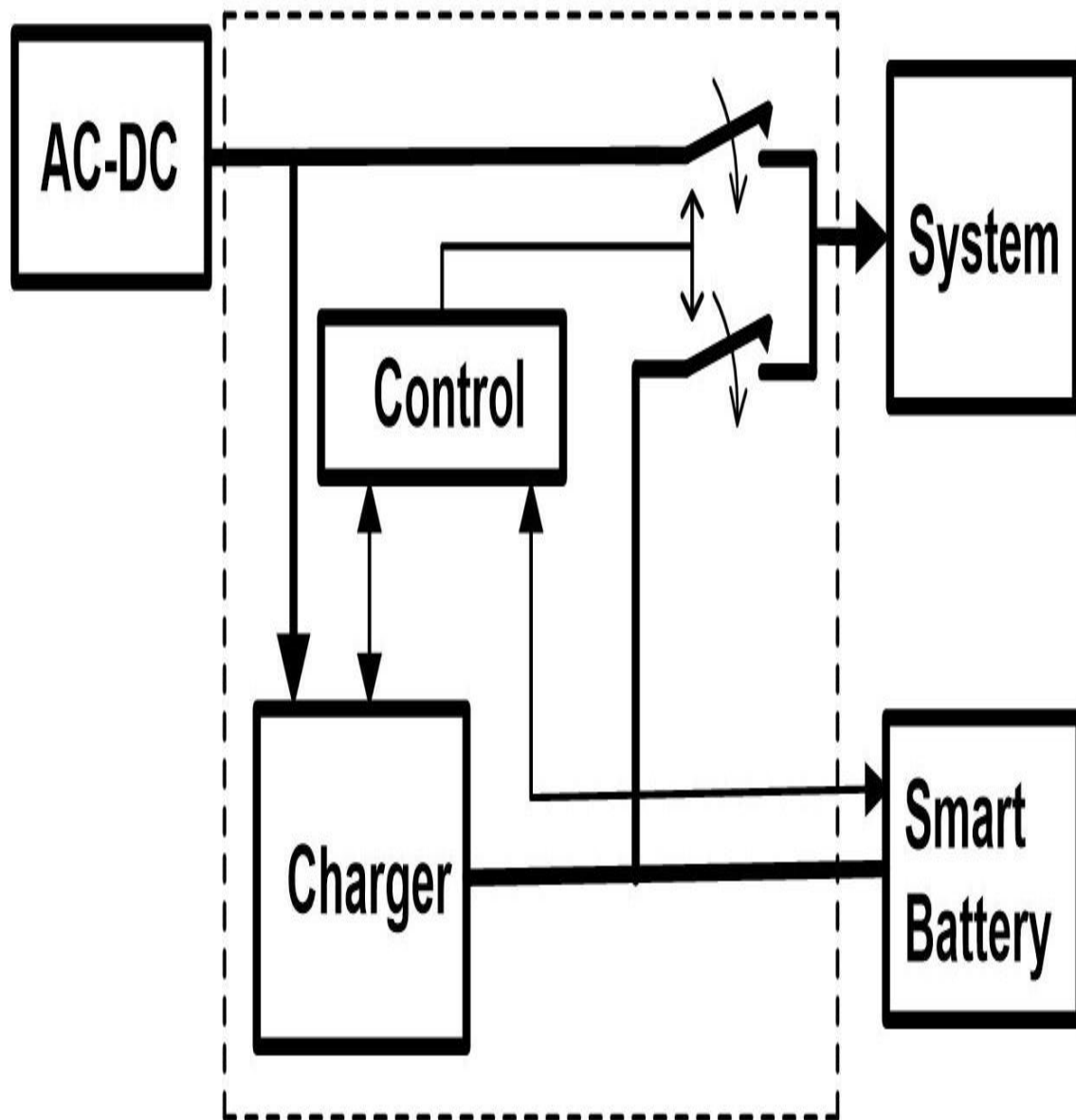


Figure 5-24. Typical Smart Battery with Charger Configuration

The above is a typical charging and battery system that would be found in a laptop computer. The system power can be provided from the AC-DC converter or the smart battery, with selection controlled by the presence of the AC-DC supply. A charger and controller provide the needed charging profile for the smart battery. All items within the dashed lines are available as a power-charger management SOC. The two switches in the power path are usually external power MOSFETs driven by the SOC.

Regulatory and Safety Issues for Batteries

In order to sell and ship devices with Li-Ion batteries a suite of tests and regulatory certifications are needed, depending on the countries involved (See: UN/DOT 38.3, IEC/UL 62133, UL20540, UL1642). The safety certifications for Li-ion battery packs in a consumer product can be costly to obtain and for low volume products may be prohibitive. Several alternatives are available. Battery pack vendors offer multiple standard configurations of 18650 cells that are already tested and safety certified. Also, a NiMH battery pack has fewer regulatory hurdles to clear and may be a possibility.

All batteries have safety concerns that need to be addressed, and those are often unique to the chemistry. Operational temperatures, limitations on charging/discharging rates, room for physical expansion within the final application, storage and shipping temperatures are all questions needing to be asked, specific to the cells being used.

The most common industry strategy is to specify a battery pack configuration and engage a specialist manufacturer of battery packs to provide an assembled product. These organizations can provide PCB circuits with safety cutoffs, charge balancing, and other smart battery features within the battery pack. As well, most of these organizations can get their products through the safety certification maze.

Other Energy Storage & Access Methods

This book is dedicated to practical design solutions. Consequently, battery storage is presented as the foremost energy storage solution for portable electronics. Other methods are often proposed, and, a brief look at these methods can be useful:

Super capacitors are often mentioned as a battery replacement, but technology vs. price point becomes the issue.

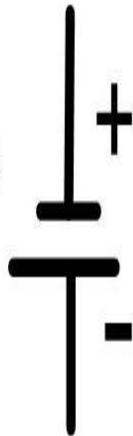
AA Battery

Alkaline

250 mA for 7.5 Hours

$Q = 7,000$ Coulombs

Price = ~\$1 USD



Super Capacitor

$C = 4000$ F @2.5 Vmax

$Q = 7,000$ Coulombs @ 1.75V

$Q = 10,000$ Coulomb @ 2.5V

Price = ~\$300 USD



Figure 5-25. Capacitor Storage vs. AA Battery

A typical alkaline AA battery costs under \$1 USD, and provides 250mA for 7.5 hours, or a total charge transfer of ~7,000 coulombs. Present (2020) technology super capacitors can be found with 4000 farads of capacitance, and a maximum voltage of 2.5V. Charging the capacitor to 1.75V, will provided the 7,000 coulombs, if discharging to 0 volts. Up to 10,000 coulombs can be stored at the maximum capacitor voltage of 2.5V. Long term reliability and how quickly the device self discharges are also an issue. The capacitor price is ~\$300 USD, and is physically much bigger than the AA battery. Generally, the lowest cost path to a good solution prevails. Here the super capacitor loses on size, cost, and reliability issues.

The idea of using a miniature hydrogen fuel cell has also been suggested.

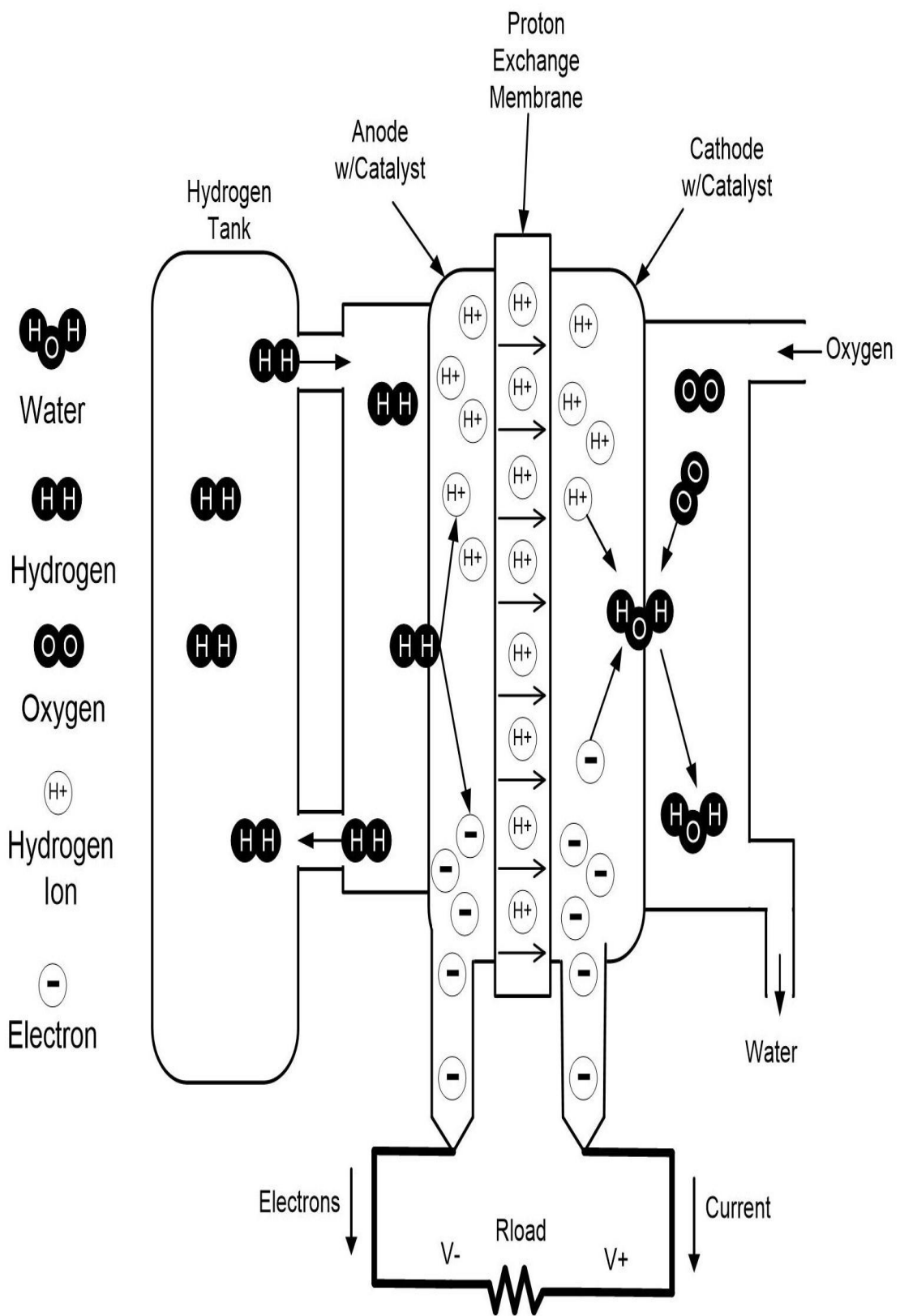


Figure 5-26. Hydrogen Fuel Cell

A hydrogen fuel cell is an electrochemical system that uses a catalyst to create Hydrogen ions, which combine with Oxygen to create a waste byproduct of water. The hydrogen ions travel across a proton exchange membrane, while the stripped electrons cause current to flow. While oxygen can be taken from the environment, in order to have a functional system, some tank or cartridge with Hydrogen is needed. Large scale systems exist, telecom systems have embraced them as backup power systems, and a limited number of cars and buses use the technology. Being able to rapidly refill a tank instead of the time taken to charge a battery is a big plus for long distance mobile vehicles.

Small scale electronics running off of fuels cells has been floated as an idea, but as of 2020 there are no small form factor devices available for commercial use. Experimental and demonstration devices do exist under 100W output, but fuels cells are better suited to higher power systems.

Flow batteries are seen as a useful device in large scale installations, but not small form factor electronics:

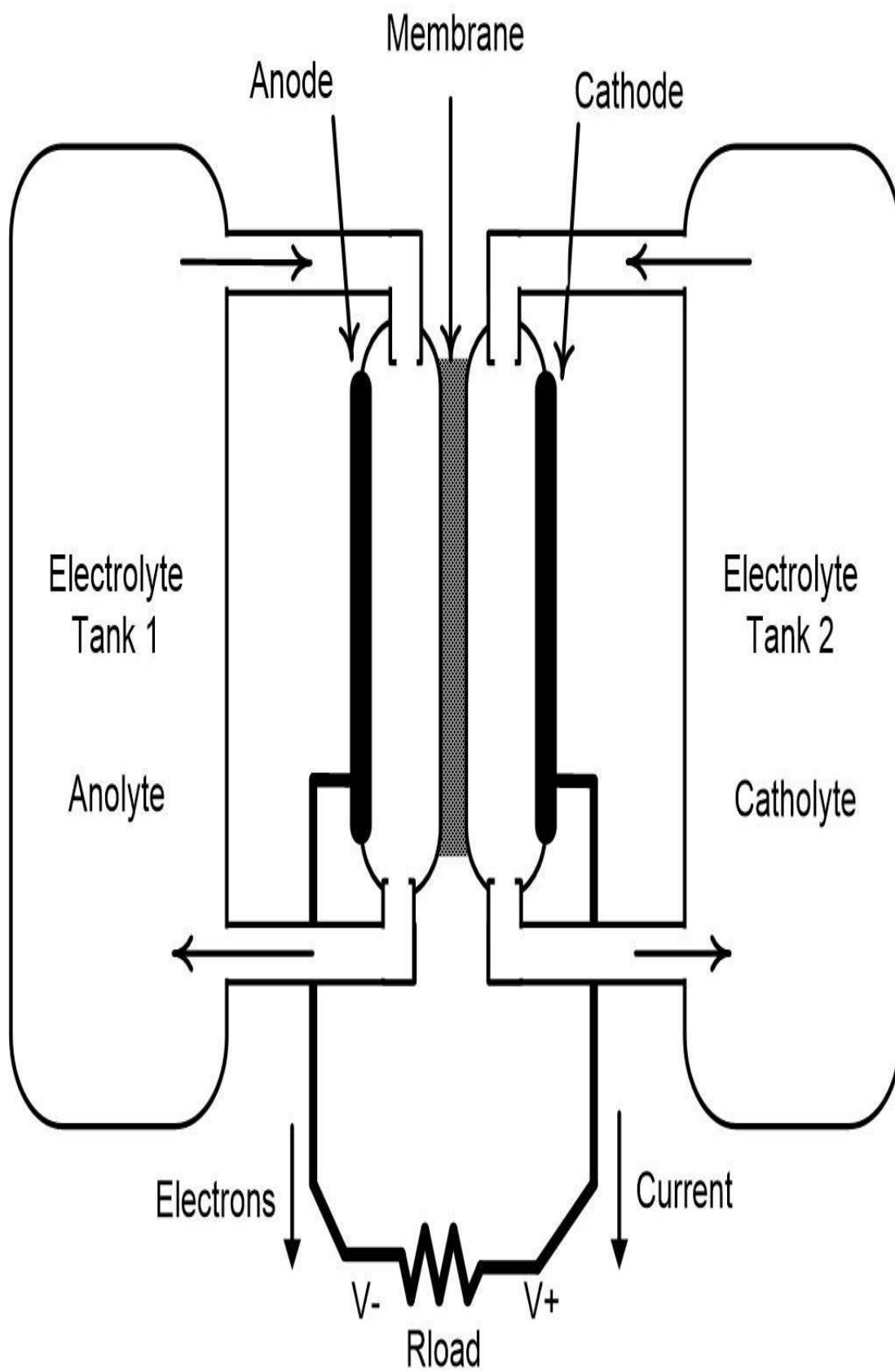


Figure 5-27. Flow Battery

The flow battery uses two electrodes, like a conventional battery, but relies on a flowing fluid electrolyte which can be pumped between the electrodes. Depending on the strategy and chemistry used, a flow battery can have a single electrolyte or two electrolytes (anolyte and catholyte) that are unique to each electrode. The amount of energy stored is linked to electrolyte volume. The maximum current is linked to the electrode size. Flow batteries have long life projections, with expected life of 20-30 years, and the ability to scale up to mega-watt-hour capacity. This makes them suitable to utility grid and enterprise scale applications. Industry has shown a strong interest in flow batteries as energy storage for solar and wind energy systems. Since they do involve tanks, pumps, plumbing, and cooling systems, these are large scale systems that are not applicable to small form factor electronics.

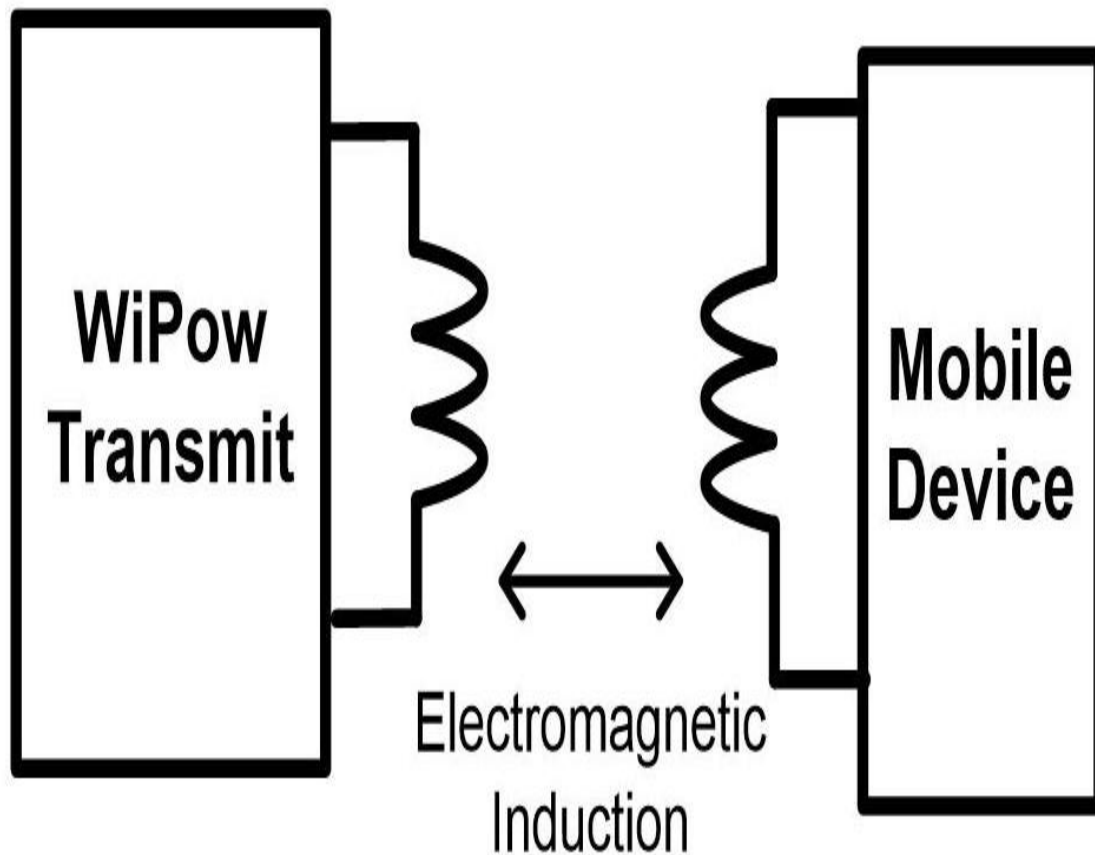


Figure 5-28. Wireless Power (WiPow)

WiPow is most commonly defined as transferring electromagnetic energy between devices without wires. WiPow can be broadly divided into several overlapping considerations – Close proximity battery charging, energy over distance, and wireless power without batteries.

Close proximity wireless charging of batteries has been widely adopted in low power (under 15 W) handheld electronics with good success. Most smart phones now include wireless charging capability. High power WiPow charging for bigger devices, such as electric cars (3-7KW), mobile medical devices (300W), robotics, etc, has also been done, with efficiency above 80%. Widespread adoption is hindered by cost. The power electronics and coils for a 300W system costs \$100-\$300 to implement, and a system for

quickly recharging an electric car sell for over \$5000. Power cords are less elegant but a whole lot cheaper

Energy over distance has problems due to transfer efficiency. It is important to recognize here that energy, and not information is transferred. A wireless signal can easily transfer data, while losing 60 dB of the signal amplitude between transmitter and receiver. Even with that signal loss the data can be recovered with some signal processing. However, a 60dB signal loss implies transmitting a million watts to recover 1 watt at the receiver. Devices closer together and using highly directional antennas improves efficiency, but that limits device mobility.

Wireless power without batteries only works if energy coupling is reliable and not interrupted. Except for controlled situations within a machine or devices tightly linked, generally it doesn't work. Energy storage is needed to get through interruptions.

Solid State Batteries hold promise for improvements in energy density (Figure 5-?).

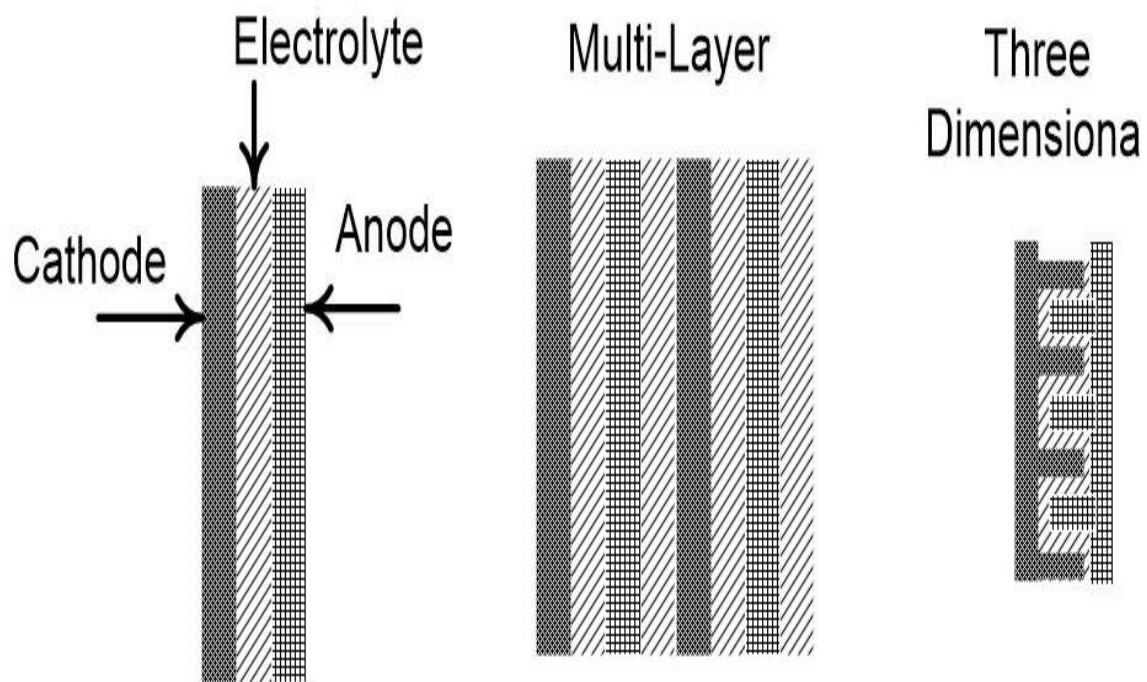


Figure 5-29. Solid State Battery

The concept of a solid state battery (SSB) can involve several things:

- Solid Electrolyte – The electrolyte is not a gel, paste or liquid.
- Lithography Fabrication Methods - Using manufacturing methods that allow better control over creation of anode, cathode and electrolyte. Chemical vapor deposition and other lithography techniques of the semiconductor fabrication world are starting to be used.
- Pure Lithium Anode - creation a pure lithium metal anode instead of the lithium compounds presently in use.

All of these items create a higher energy density battery. SSBs are presently limited to small low power devices and expensive limited quantities. Many SSB research efforts are ongoing in industry. At the moment, most sources are promising results “5 years” in the future. Issues remain with processes and methods, scaling production to volume and making the devices cost competitive with existing technology.

Essential Concepts – Conclusions and summary

The important ideas and items of this chapter are:

- Battery specifications are not standardized, and are inconsistent between vendors. Need to read the details and compare the test conditions.
- Deciding between single use or rechargeable devices.
- Estimating or measuring the system power requirements.
- Fitting the battery discharge profile to the power requirements of the system.
- Battery types by chemistry, useful features, pros and cons.
- Equivalent circuits for batteries

- Battery capacity changes with load and is not a fixed value.
- Battery capacity changes with temperature, cold batteries have lower capacity
- Performance summaries for commonly available batteries
- Load balancing of high current battery connections
- Safety cutoff circuits and charge balancing for battery packs
- Constant current charging and charging profiles by device chemistry
- Alternative energy methods

After reading this chapter, the reader has the knowledge to:

- pick the battery chemistry,
- determine needed capacity,
- pick a suitable device package,
- define the serial/parallel configuration,
- define a charging strategy, and
- define smart battery features where needed.

Manufacturers of battery packs can provide the most recent specification for cells they have available and build the device specified. Semiconductor manufacturers can provide charging controllers suitable to design a charging circuit that can be fitted to the parameters of the battery pack.

Further Reading

“The TAB Battery Book: An In-Depth Guide to Construction, Design and Use”, 2011, Michael Root, McGraw Hill, ISBN_ 978-0-07-173990-0

“Batteries in a Portable World” 3rd Edition, 2011, I. Buchmann, Cadex Electronics, ISBN 0-9682118-2-8

“Simple Pspice Models let you Simulate Common Battery Types” S. C. Hageman, Electronic Design News, vol. 38, pp. 117 – 129, 1993

“Solid-State Batteries Inch their way Toward Commercialization” M. Reisch, November 20, 2017, Volume 95, Issue 46, Chemical and Engineering News, Retrieved from: <https://cen.acs.org/articles/95/i46/Solid-state-batteries-inch-way.html>

“Key Issues Hindering a Practical Lithium-Metal Anode” C Fang, X Wang, YS Meng, Trends in Chemistry, vol. 1, Issue 2, P152-158 Retrieved from: [https://www.cell.com/trends/chemistry/fulltext/S2589-5974\(19\)30028-0](https://www.cell.com/trends/chemistry/fulltext/S2589-5974(19)30028-0)

“Power Consumption Analysis of Bluetooth Low Energy, ZigBee and ANT Sensor Nodes in a Cyclic Sleep Scenario” A. Dementyev, S. Hodges, S Taylor, J Smith, Department of Electrical Engineering, University of Washington, Seattle, Retrieved from: <https://semiwiki.com/wp-content/uploads/2016/08/IWS20201320wireless20power20consumption.pdf>

“An Analysis of Power Consumption in a Smart Phone” A Carroll, G Heiser University of New South Wales, Retrieved from: https://www.usenix.org/legacy/event/atc10/tech/full_papers/Carroll.pdf

“Flow Batteries: Energy Storage Option for a Variety of Uses” T Scroggin-Wicker, K McInerney, March 1, 2020, Power Magazine, Retrieved from: <https://www.powermag.com/flow-batteries-energy-storage-option-for-a-variety-of-uses/>

“Battery University” I Buchmann, Cadex Electronics, Retrieved from: <https://batteryuniversity.com/>

“Power and Battery Technical Resources for Design Engineers – PowerStream” Retrieved from: <https://www.powerstream.com/tech.html>

“Engineering Guidelines for Designing Battery Packs – PowerStream” Retrieved from: <https://www.powerstream.com/BPD.htm>

“Custom Battery Packs, Technical Engineering Guide Epec Technologies” Retrieved from: <https://www.epectec.com/guides/>

“Battery Management Deep Dive On-Demand Technical Training” Texas Instruments, Retrieved from: <https://training.ti.com/battery-management-deep-dive-technical-training>

Chapter 6. Electromagnetic Interference and Electrostatic Discharge

A NOTE FOR EARLY RELEASE READERS

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

This will be the 6th chapter of the final book.

If you have comments about how we might improve the content and/or examples in this book, or if you notice missing material within this chapter, please reach out to the editor at arufino@oreilly.com.

Introduction

Electromagnetic Interference (EMI) deals with unwanted signal coupling and other secondary effects within or between electronic systems. The terms; Electromagnetic Compatibility (EMC) and Signal Integrity are often used as generalized terms to cover most of these topics. EMI and Radio Frequency Interference (RFI) are often used interchangeably, and herein are considered the same thing. Historically, “RFI” has been in use since the early era of analog radios. In this book, the terms EMI, RFI, Crosstalk, and Noise are used interchangeably. Electrostatic Discharge (ESD) is included in this chapter as a special topic under the umbrella of “unwanted” signals. Many circuits discussed here affect both EMI and ESD performance.

Any digital electronic system creates EMI, with the noise magnitude determined by clocking rates, number of gates, and quantity of many other switching devices. Designers should recognize that the environment is awash in electronic noise, due to active electronics, radio devices, solar noise, grid power systems and countless others. Designers need to limit the magnitude of the EMI created by their designs, and protect their electronics from the noise of the outside environment.

Within an electronic system, noise generated can also be a detriment inside the same system. Sometimes, internal noise can reduce accuracy, degrade performance or make a device nonfunctional.

ESD events also need to be dealt with as part of system protection. For consumer products, ESD protection must be sufficient to avoid functional destruction. For high reliability designs, like military, industrial and medical devices, the protection must maintain proper functionality even in the presence of ESD events.

Preliminary Ideas

Many ideas presented here can be implemented with little or no cost. Some require addition of low cost items like ferrite beads (FB), resistors, and capacitors. Including EMI reduction measures into a design is important for getting regulatory clearance. Many consumer products have a FB molded onto the power cord, due to failed emissions testing. That FB was added to the power cord as a last minute fix. That could have been avoided.

Generally, systems that require comprehensive EMI measures all switch currents at high speed. Switching power supplies are noise generators. All digital devices generate EMI. Anything with a clock over 100 MHz commonly needs noise reduction techniques in place. Small devices like hand held battery powered devices with clocks under 1 MHz generally don't need noise suppression. That leaves a large grey zone in the middle where noise efforts may, or may not, be needed.

Including many of the low/zero cost EMI reduction techniques of this chapter will help keep devices within regulatory limits. Anything with

conductive connections or contacts outside of an enclosure requires ESD protection.

From outside the box, EMI gets evaluated in four areas (Figure 6-1).

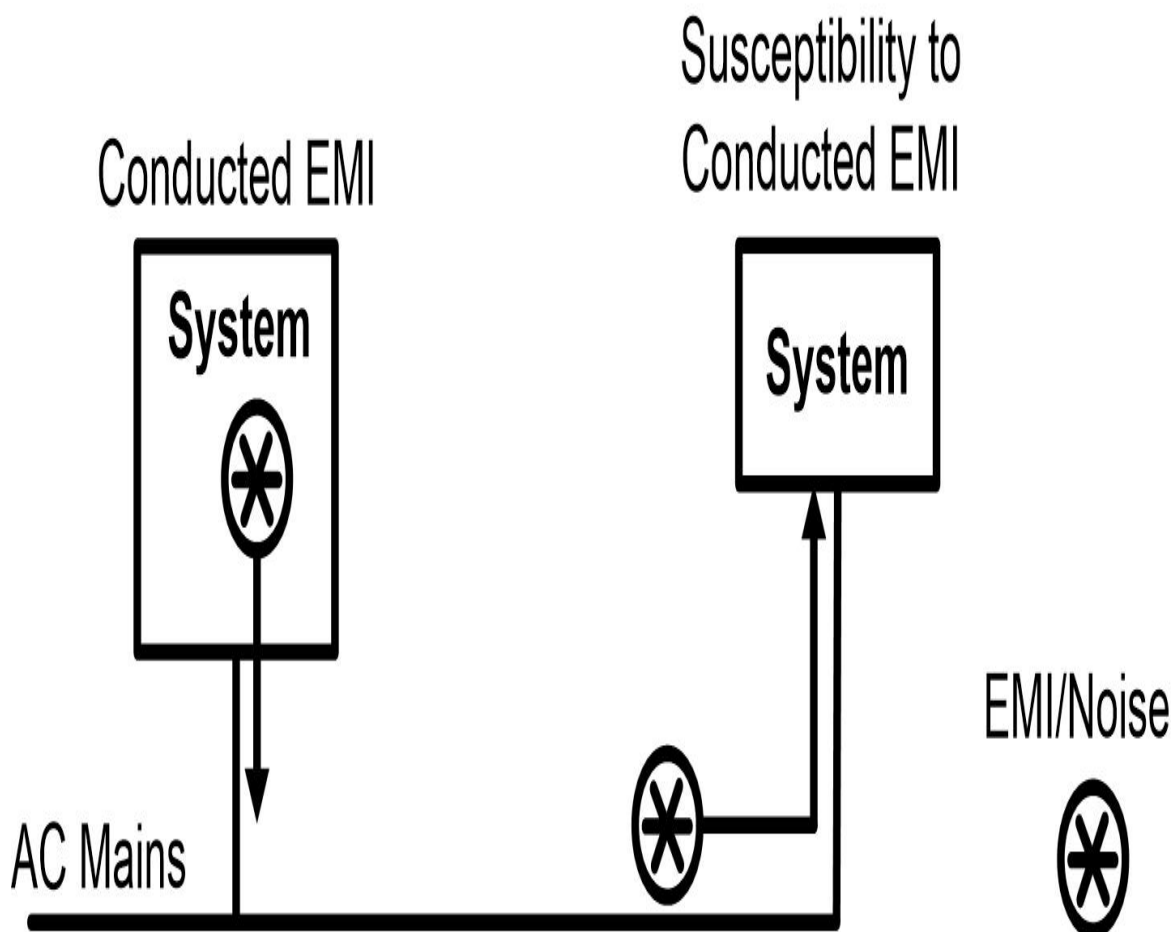
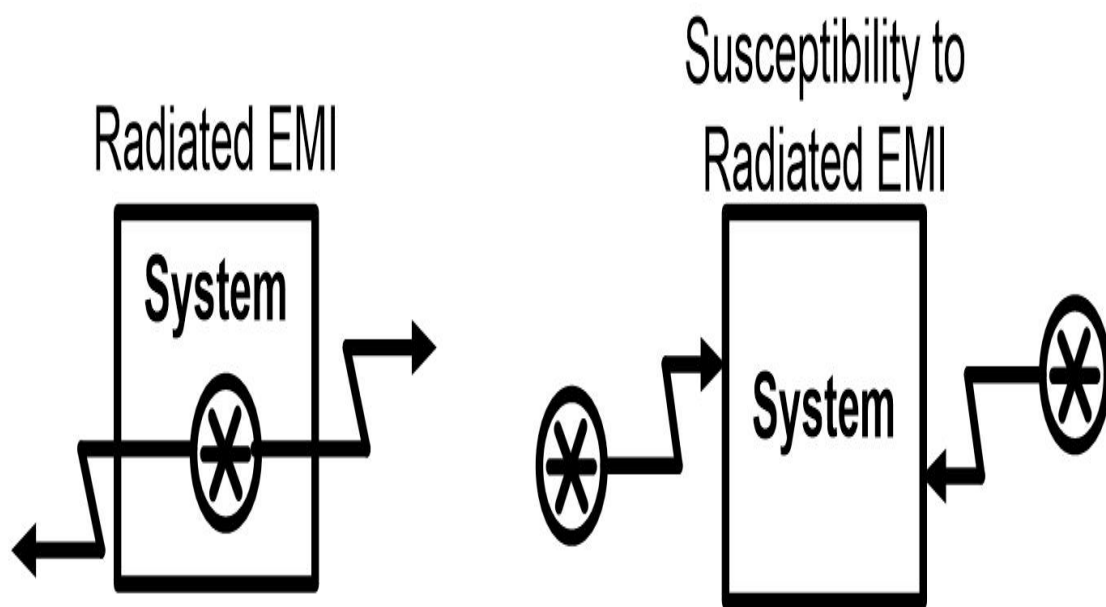


Figure 6-1. EMI Four Avenues at the Top Level

Radiated EMI is measured from the perspective of externally observed noise. Does the system under test generate so much EMI that it affects the performance of other devices outside the system?

Susceptibility to Radiated EMI is the opposite side of the problem. For defined amounts of EMI external to the system, does it affect the system performance?

Conducted EMI deals with noise, created by the system, that is injected onto the AC power lines. This manifests itself as current transients, or high frequency voltage transients, that the AC power mains can't suppress.

Susceptibility to Conducted EMI deals with a system's capability to properly function while attached to noisy, non ideal AC power mains.

Depending on the product, industry standards or regulatory compliance requirements have been defined to test these EMI scenarios. Medical, aerospace-aviation and military devices have well defined mandatory testing for product qualification. Other industries and government regulatory bodies have developed requirements unique to their sectors.

Internal to a system, dealing with EMI encompasses several areas: radiated noise created, noise coupling within the system, and immunity to external noise.

Electronic noise can come from sources, other than EMI. Inherent/Intrinsic noise is due to the characteristics of system components and semiconductors (Figure 6-2).

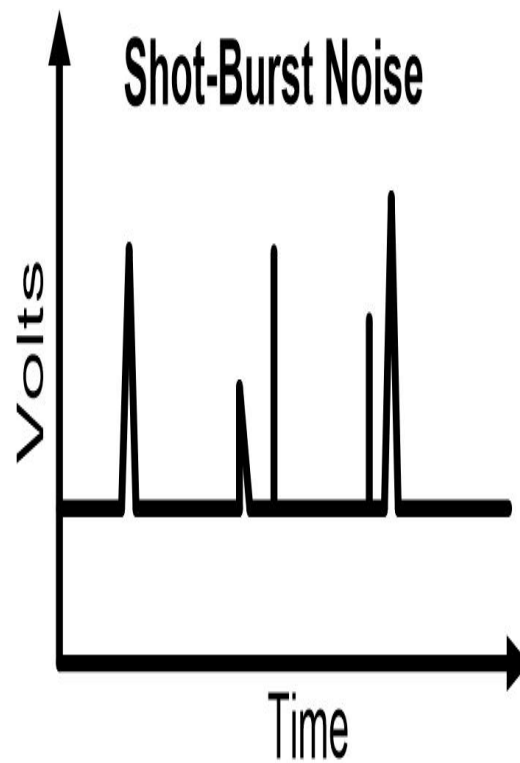
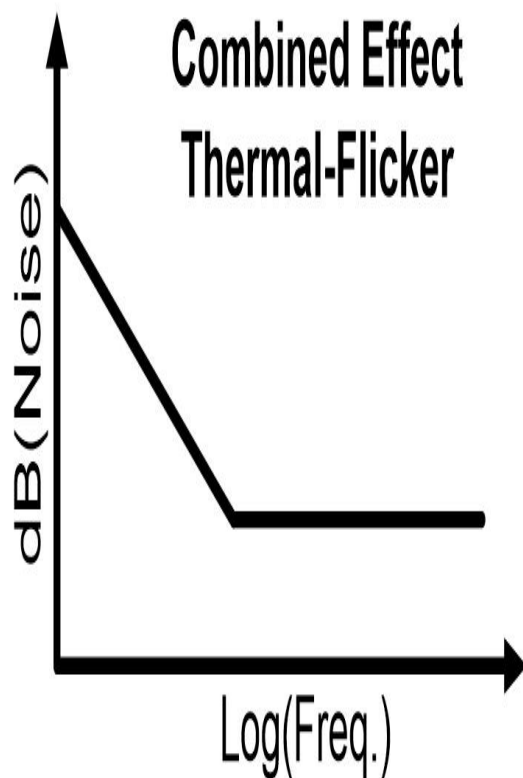
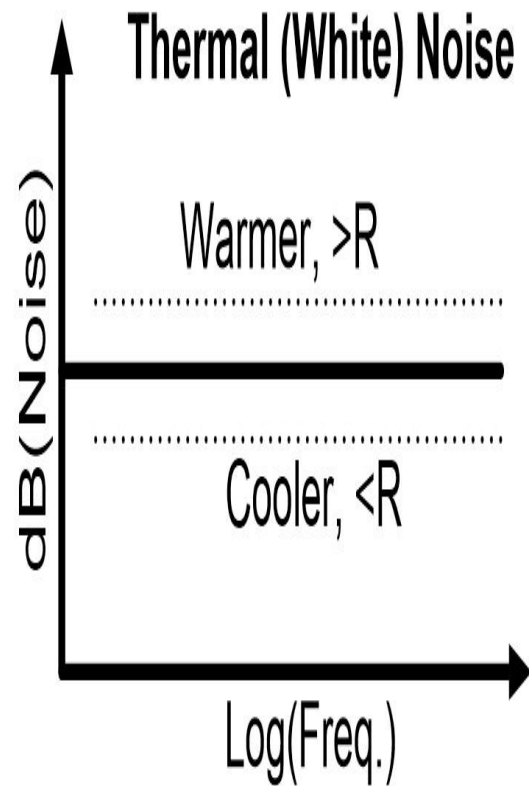
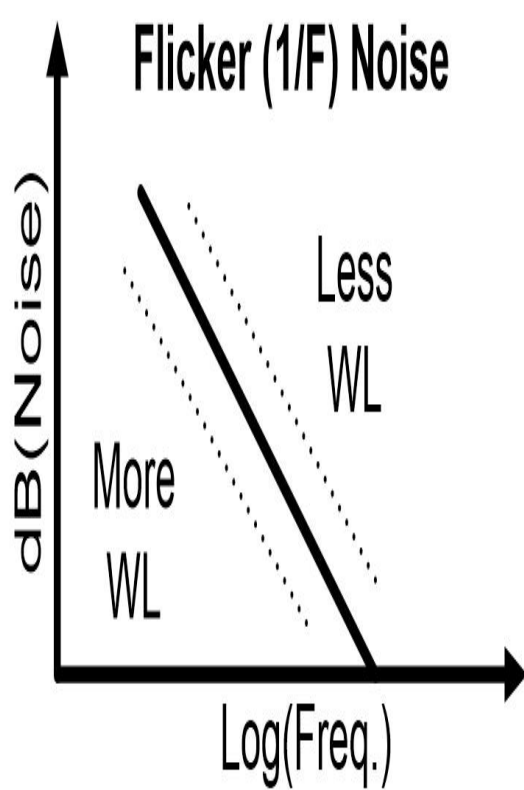


Figure 6-2. Intrinsic Noise Sources

Flicker noise (also known as $1/f$ noise) is a low frequency component primarily from CMOS transistors, related to the non ideal characteristics' of transistors. Magnitude of the noise is proportional to the size of the transistor, with larger devices creating less noise.

Thermal noise (also known as Johnson-Nyquist noise) is wideband noise (aka white noise) with no spectral weighting (flat over frequency) due to thermal effects within resistors. Higher temperatures or bigger resistance creates more noise.

Frequently an amplifier will exhibit a combination of thermal and flicker noise. A low frequency ramp down to a flat horizontal noise floor is commonly seen.

Shot noise (also known as Poisson noise) and Burst noise (also known as popcorn noise) are both due to the transition of discrete charges (single electrons or small numbers of electrons). Burst noise is semiconductor related and caused by the trapping and releasing of electrons. Shot noise is associated with the current in a conductor, and is seen at extremely low currents where singular electrons are significant proportion of the current.

Intrinsic noise sources are issues of concern within AFE devices and analog IC design. System level designers will generally not need to deal with them, but should be aware of their existence. The focus here is EMI, both in/out of systems, and how to mitigate it.

General Strategy Dealing with EMI

The approach used here builds an awareness of noise issues that will be encountered, and illustrates design methods to manage those issues. System noise needs to be limited in order to remain functional and accurate. Noise needs to be restricted outside your system in order to be sold commercially.

It is quicker and cheaper to deal with noise issues as part of the design process, rather than trying to fix problems in a finalized design. To properly

mitigate noise issues, designs need low impedance grounding, properly protected ports, and minimization of radiated noise. In addition, a functional tolerance for both internal and externally applied EMI is necessary. In many cases, EMI limiting enclosures (Faraday cages) are also needed to meet regulatory restrictions.

Ignoring EMC/ESD issues invariably results in one thing: Devices failing regulatory testing. That quickly leads to a haphazard attempt to duct tape in filters, ferrites and ESD protection as the product readies for production. That approach costs more than implementing EMC-ESD measures within the design.

The approach outlined here is one of best practices, to optimize or improve the design. Quantitative analysis of radiated noise, coupled with an analysis of noise susceptibility, is theoretically possible, but neither practical nor expedient. In a complex system, it is not one noise source and one victim of noise, but hundreds of sources and victims. To do a quantitative analysis is extremely time consuming. Instead, an efficient approach is to recognize and resolve problematic areas, while including methods within the entire design to improve noise performance. The outcome is both robust and reliable.

Regulations and Requirements

What regulatory requirements apply depends upon the device's application and country of use. All regulations and requirement approach devices as a black box, either inflicting EMI, or measuring EMI, both radiated and conducted, while observing defined functionality.

The international standards group that defines EMI requirements is CISPR (Comité International Spécial des Perturbations Radioélectriques, loosely translated as: International Special Committee on Electrical Radio Interference). CISPR is part of the IEC. Most countries provide legal requirements that mandate CISPR standards or a similar variation thereon.

Within the USA, the FCC regulates most devices. The FDA regulates medical devices, and the DOD issues requirements for military electronics.

Depending on the product, other agencies may issue requirements for their sector.

A typical FCC restriction on radiated emissions can be found in the Code of Federal Regulations Title 47 (Telecommunications), Volume 1, Chapter 1 (FCC), Subchapter A, Part 15, Subpart B (Unintentional Radiators) paragraph 109 (Radiated Emissions Limits). After interpreting the regulation it can be simplified to a “stay under the line” compliance plot (Figure 6-?).

FCC Radiated Emissions Limits, Title 47, Part 15 B

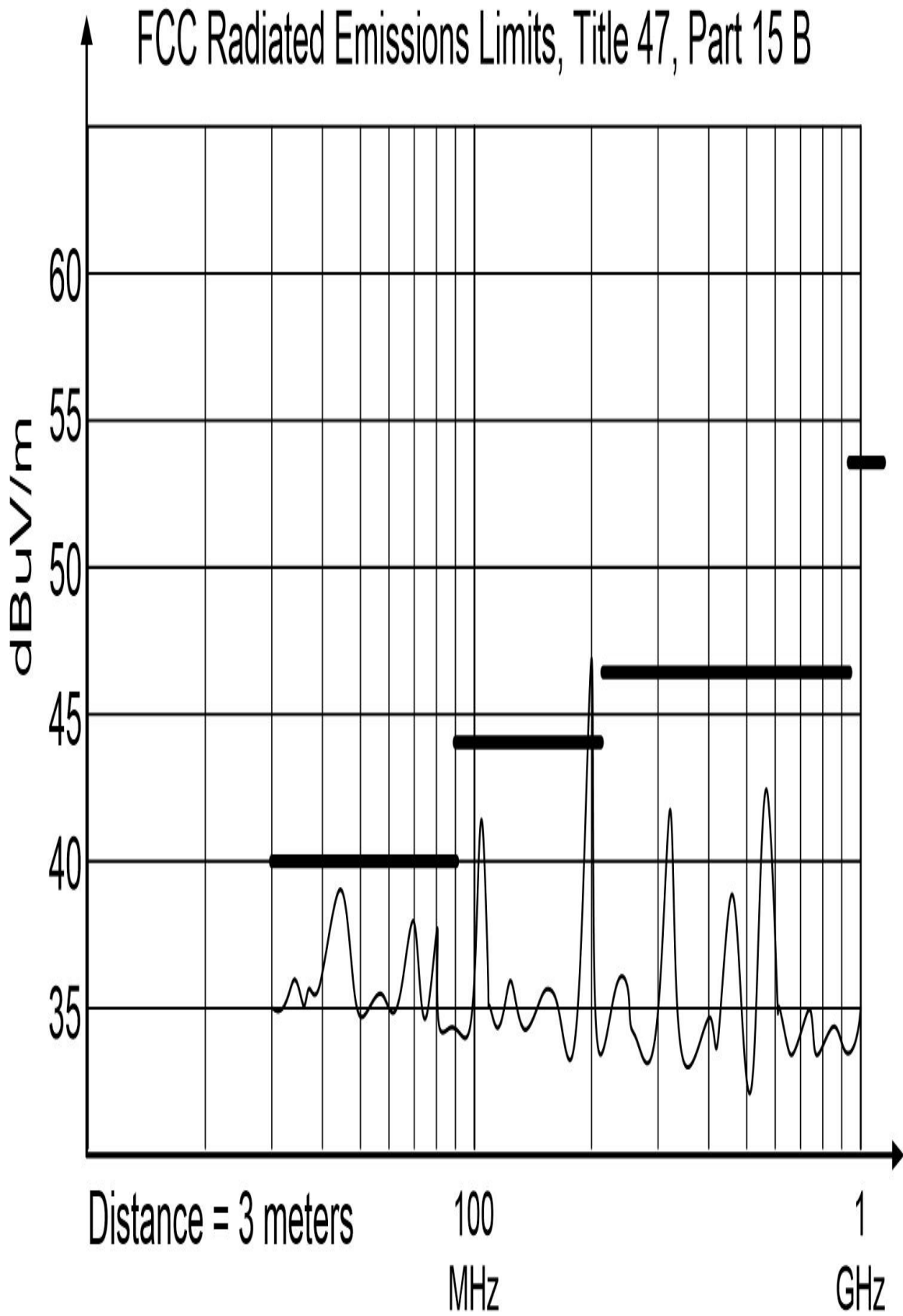


Figure 6-3. FCC Part 15B Emissions Restrictions

The horizontal bars indicate the maximum allowed radiated noise over frequency. The example shown has a violation at 200 MHz, but is otherwise compliant.

Two common ways to test for radiated emissions are using an Open Air Test Site (OATS) or within an Absorber Lined Shielded Enclosure (ALSE) screen room. The ALSE setup is similar to an OATS setup except being contained within a Faraday Cage screen room. Both yield suitable results when calibrated and properly run. An OATS setup looks like Figure 6-?.

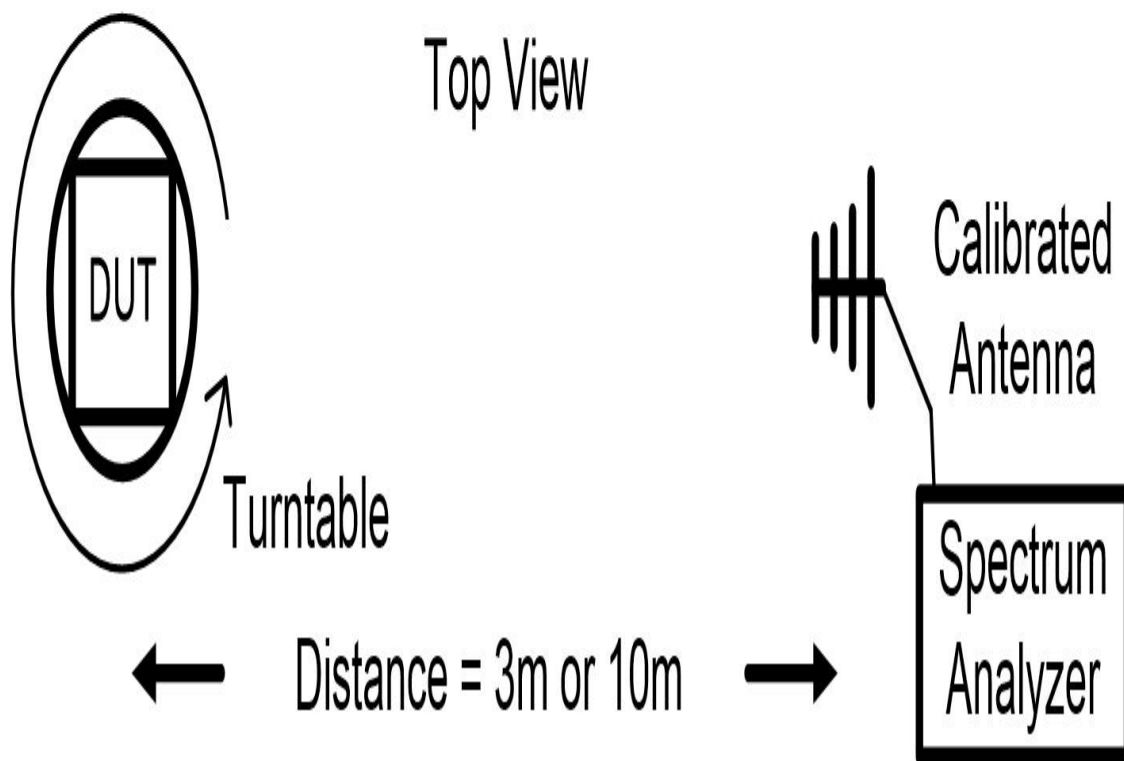
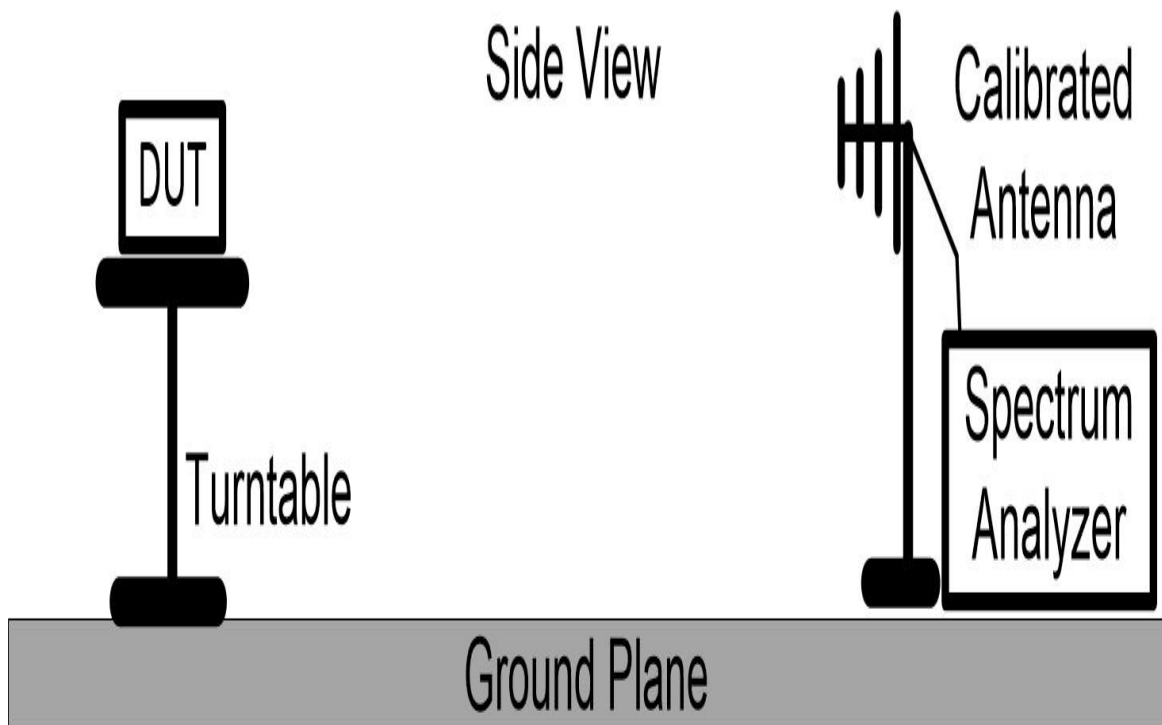


Figure 6-4. OATS Testing Setup

A calibrated antenna is used to observe radiated emissions at a fixed distance (3M or 10m as per FCC requirements) with a spectrum analyzer attached. Since emissions from the Device Under Test (DUT) can be directional, the DUT is rotated 360 degrees while scans are run.

All commercial devices in the USA need to meet either FCC Part 15A or 15B for radiated noise. The FCC (Part 68) also covers conducted emissions put back onto power lines, and other regulations are applicable. Depending on the product sector, applicable requirements vary. A full determination of what regulatory requirements need to be met should be made early in the design process.

Devices purchased by the DOD must meet Military Standards 461 (revision G, as of 2015). Items in the European Union need to meet CISPR 22 standards for EMC. Japan has its own standards body in the VCCI Council (See: vcci.jp).

Visualizations of Noise Coupling

Some lab testing will be a part of any noise analysis. This is not a book on proper test and measurement (T&M) techniques, so readers are cautioned to educate themselves on proper T&M methods to get reliable data. Much of the data mentioned here can be generated by a certified test laboratory.

One of the clearest illustrations of noise scenarios is a “Talker-Listener” model, with a transmission medium in between. Frequently this is called the “Source – Medium – Victim” model (Figure 6-?).

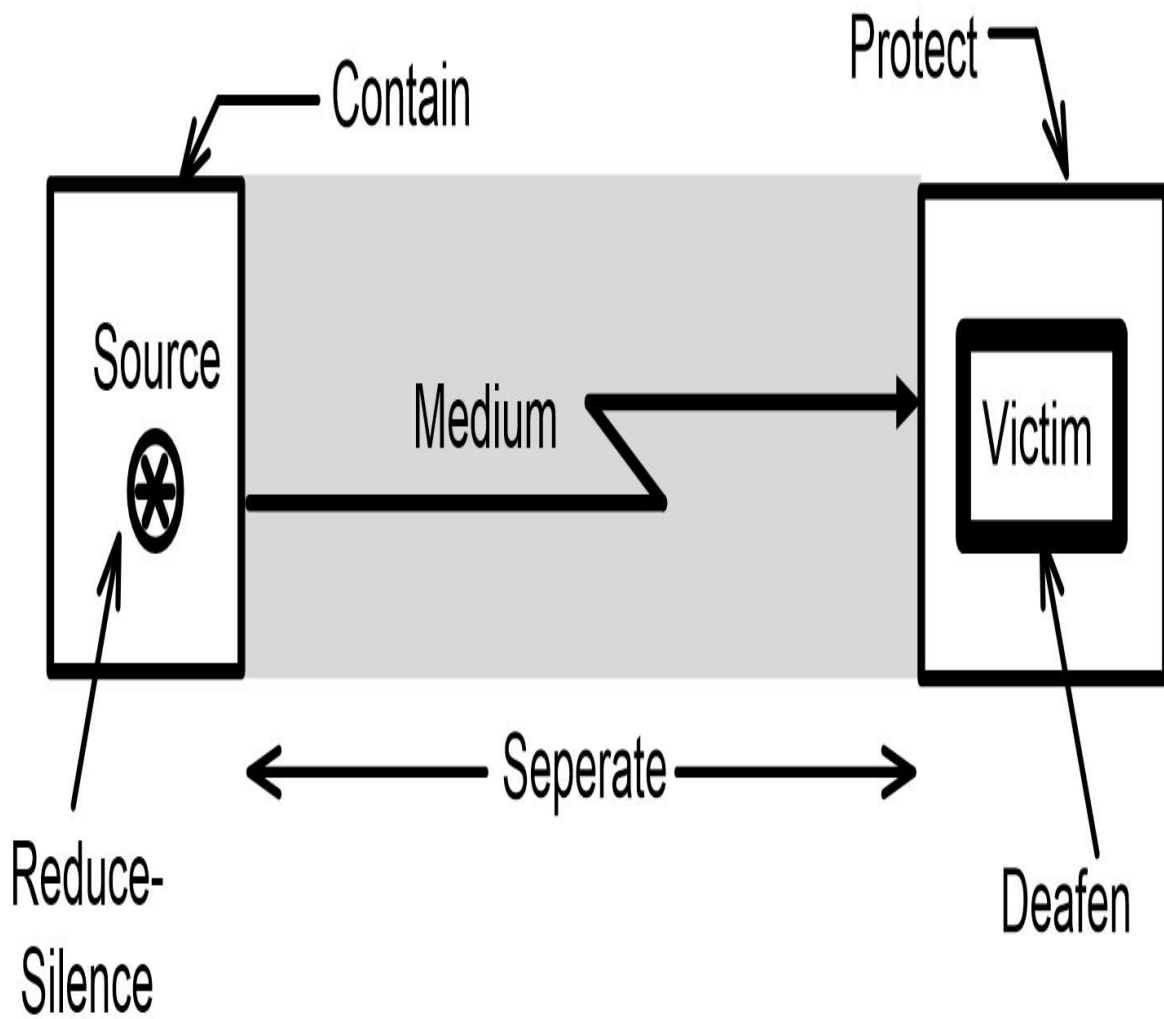
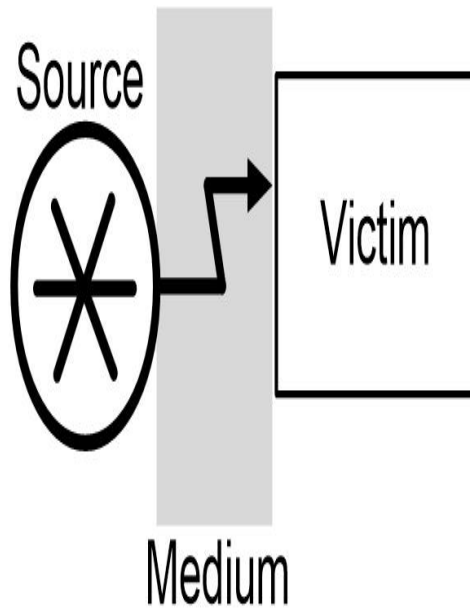


Figure 6-5. Source – Medium – Victim Noise Model

Understanding noise issues can all be fitted to this model. Namely:

Silence the Source – Techniques used to reduce the magnitude of noise generated by any given source.

Contain the Source – Techniques that don't silence the source, but methods to localize the noise generated and keep it at or near the source.

Protect the Victim – Methods such as shielding and enclosures that don't allow noise near the victim. Filtering and bandwidth limiting connection signals are also included here.

Deafen the Victim – When noise gets to the victim, but the victim is non-responsive. Circuits with common mode noise rejection, or bandwidth limiting, are less responsive, and often useful here.

Separate the Source & Victim – The most common strategy here is to physically separate with distance. Separation in frequency can be used in some cases and time interleaving, separation in time, of noisy events and noise sensitive events can also be done.

Getting useful information about noise sources and their cause is best determined in the frequency domain (Figure 6-?).

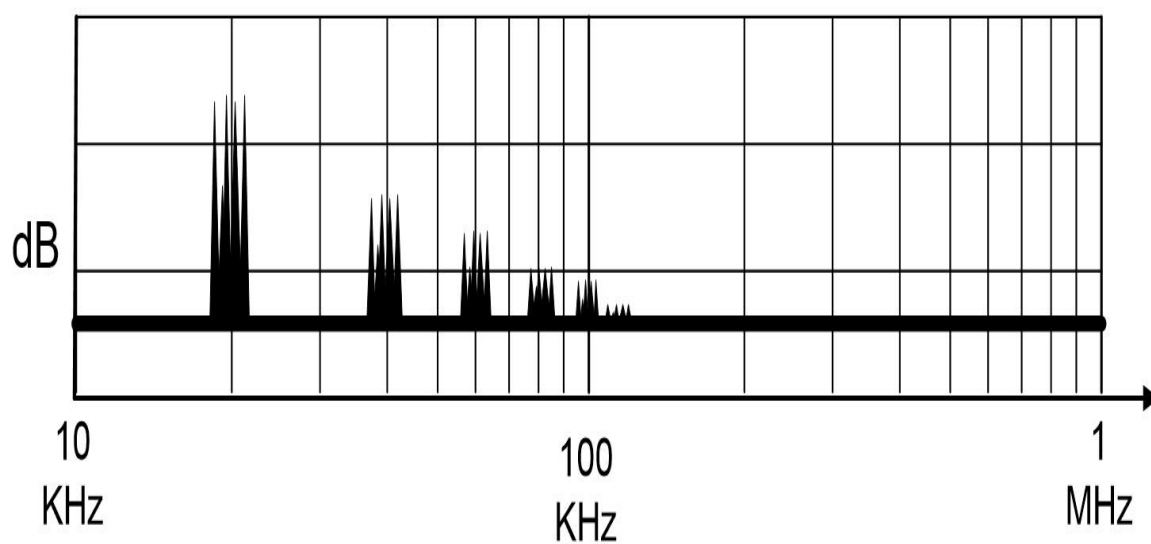
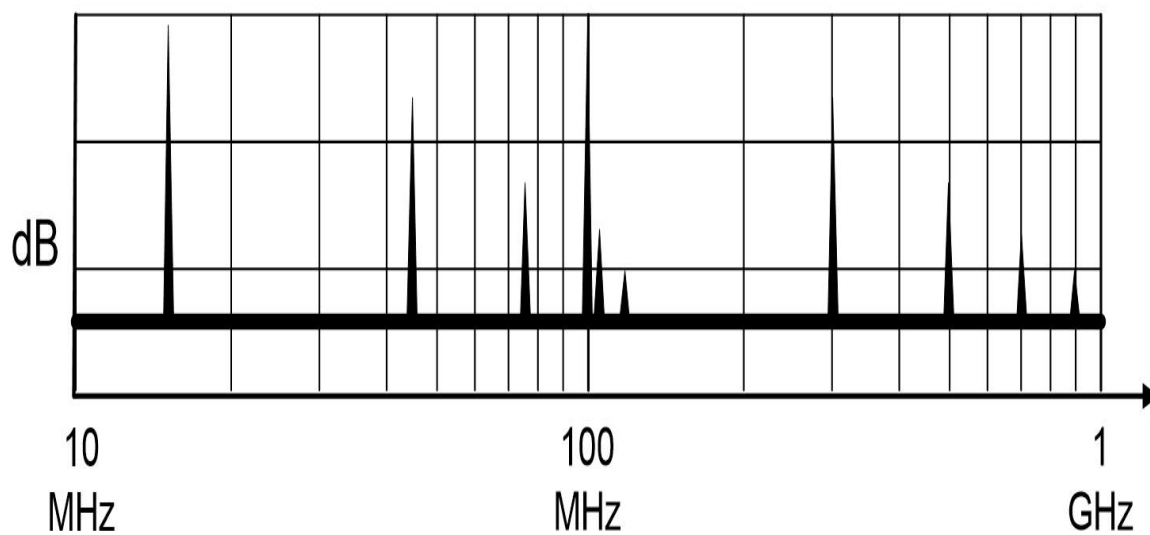
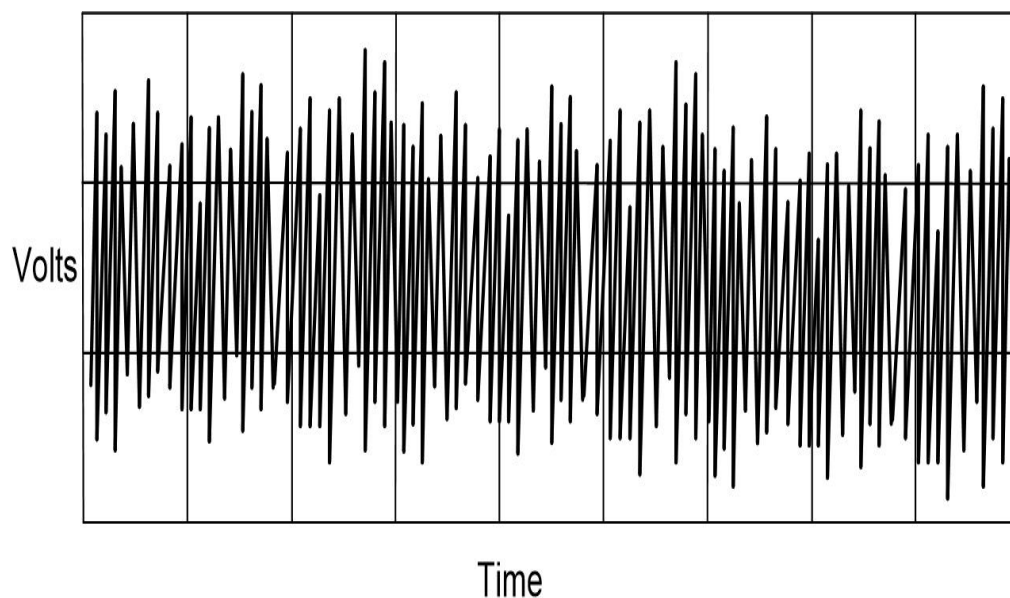


Figure 6-6. Noise in Time vs. Frequency Domain

The top plot is voltage vs. time domain, and there's not a lot of readily apparent information there. Looking at the signal spectrum quickly tells a clear story. In the frequency domain there are spikes at 15 MHz, 45 MHz and 75 MHz, which are harmonic multiples of 15 MHz. In this example, the system has a free running clock at 15 MHz which is radiating outside the device. In addition, this example shows signals at 100, 300, 500, 700, and 900 MHz. Again, another clock in the system is at 100 MHz, and its harmonics, are part of the noise problem.

At lower frequencies noise is seen at 20 KHz, and harmonics thereof. This would be typical noise due to a switching power supply operating near 20 KHz. When the frequency content is examined, there is often a very clear message of what is creating the noise observed.

When dealing with noise issues the analogy of “peeling the noise onion” is frequently used. The author prefers the analogy of an archeology dig; as you dig deeper, and more noise problems are discovered and resolved, another layer of noise issues become apparent. As an example, here's a noise profile before and after some noise suppression was added (Figure 6-?).

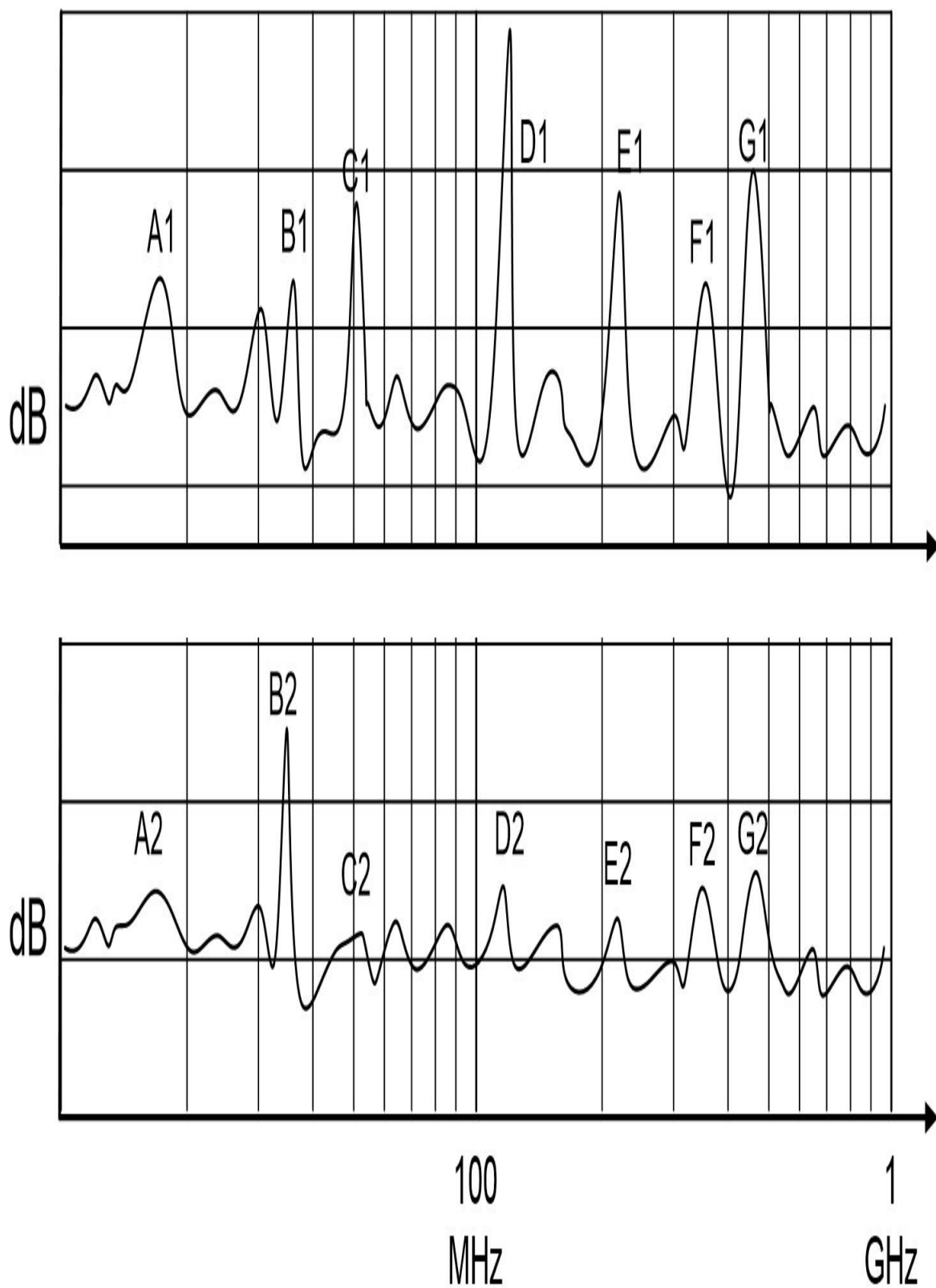


Figure 6-7. Noise Profiles as Sources are eliminated

Before - The upper plot shows the biggest noise contributor “D1” peak slightly above 100MHz.

After - The lower plot shows that the D1/D2 signal has been reduced and is no longer the dominant noise problem in the system. Most of the high frequency noise problems above 100MHz (D1, E1, F1, G1) have been reduced, and the significant noise problem is at B2 (~35 MHz).

These are typical results. As noise sources are eliminated or reduced, other noise sources become apparent.

Most EMI is created from transient switching, primarily in digital devices and power supplies. Both create large amounts of noise. Understanding how noise propagates to other devices thorough a system is useful (Figure 6-?).

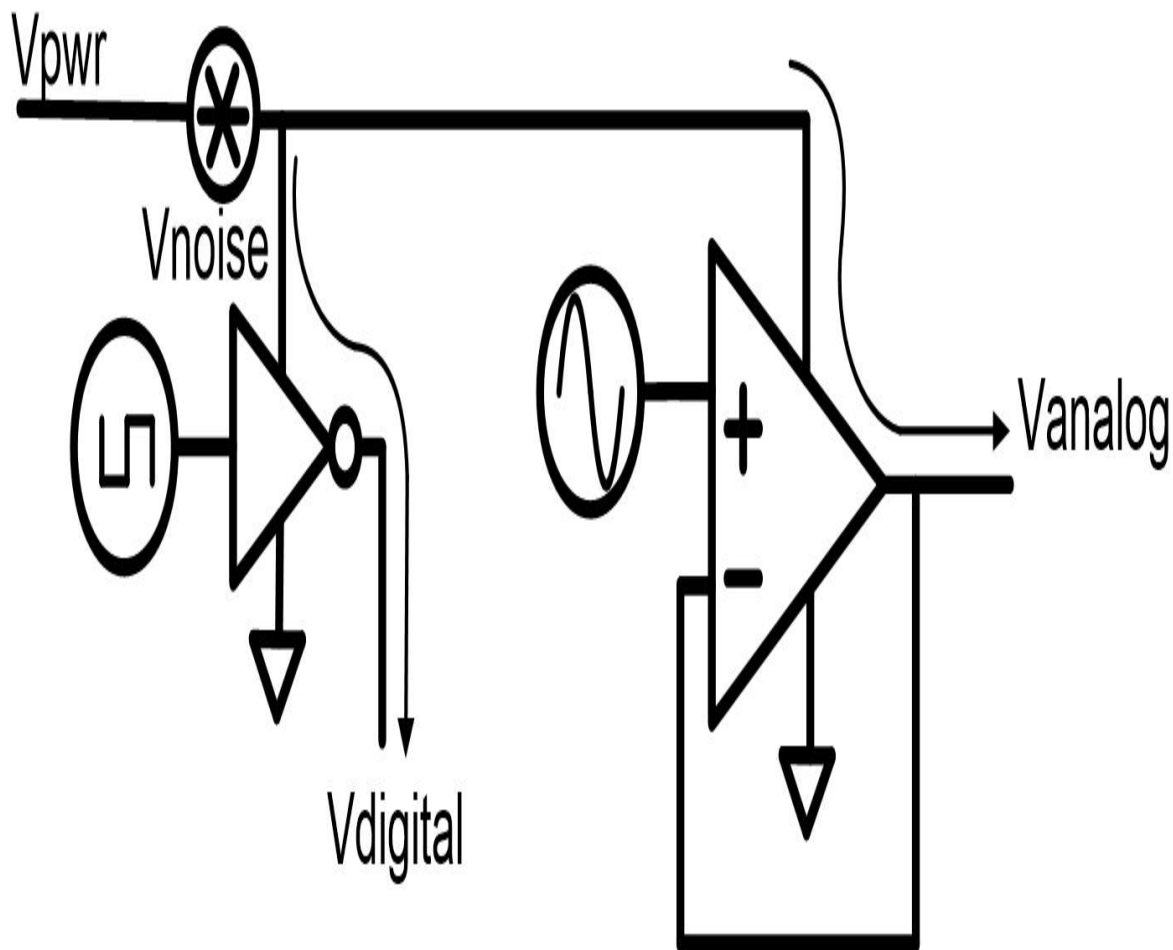


Figure 6-8. Noise Coupling through Power and Ground (P&G) paths

P&G connections are common to all circuits on a PCB. Noise on either of the P&G connections gets to all circuits. Consequently a low impedance common ground is important, and a power supply network with high frequency stability is important. However, some noise will exist on both P&G due to less than perfect power systems, connection impedance and multiple switching transients throughout the system.

For digital devices, P&G noise transfers directly to logic outputs, primarily through internal transistor connections and parasitic capacitance. For synchronous digital devices some noise is tolerable. With devices that are not synchronized, noise transfers into time domain edge jitter. P&G noise present on rising or falling digital signals transfers to time variance.

For analog and mixed signal circuits, P&G noise will transfer over to output signals or corrupt internal circuit functionality. Depending on internal circuits, noise can be directly transferred, or be actively amplified by internal circuits. Analog and digital devices directly connected to the same power supply can be especially problematic.

Coupling of noise also occurs through parasitic capacitance between connections and magnetic induction methods (Figure 6-?).

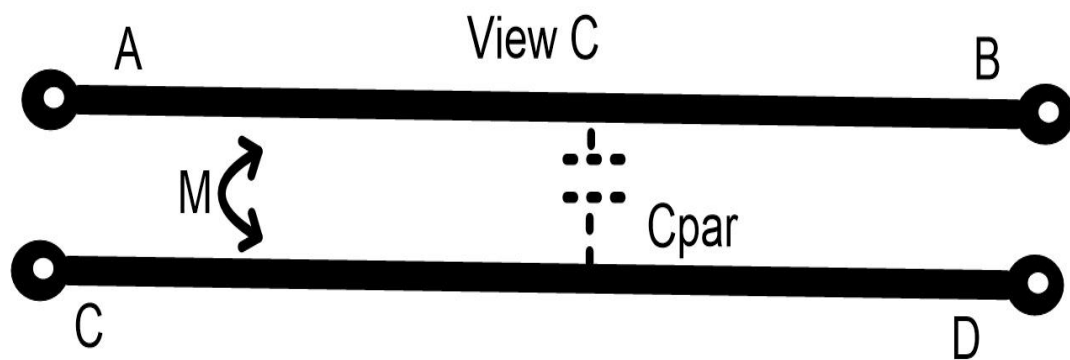
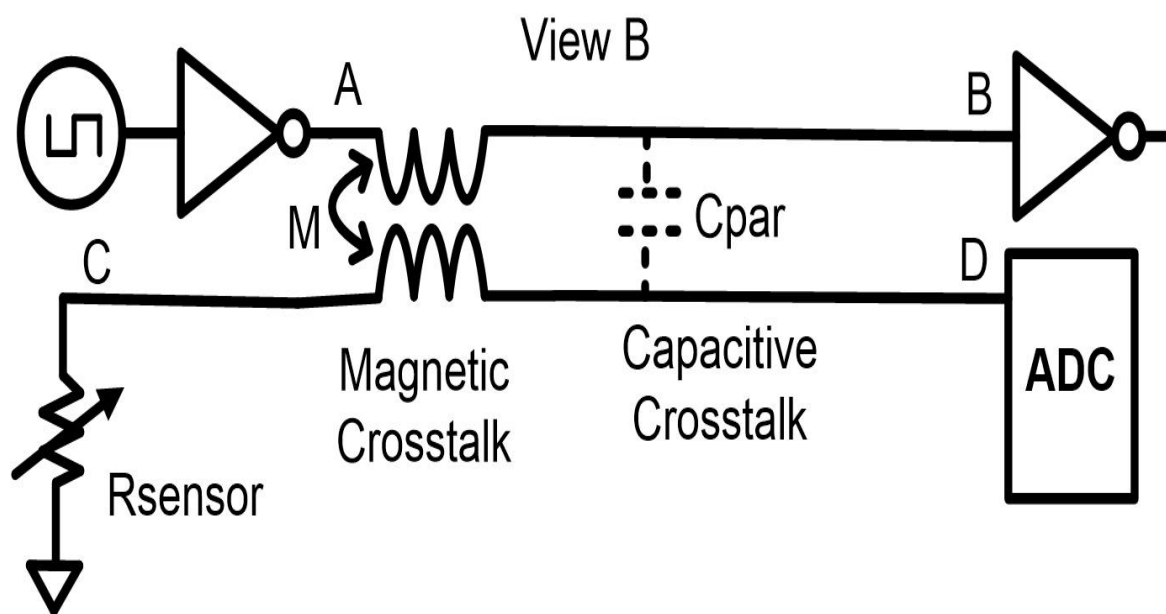
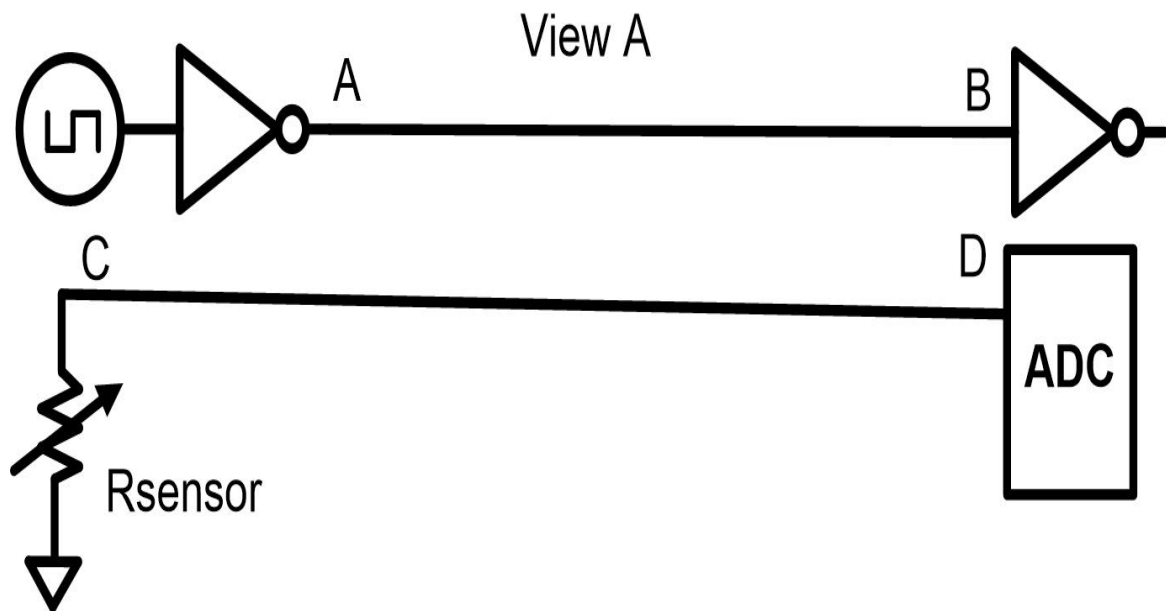


Figure 6-9. Noise Coupling through Capacitive/Magnetic (Inductive) Methods

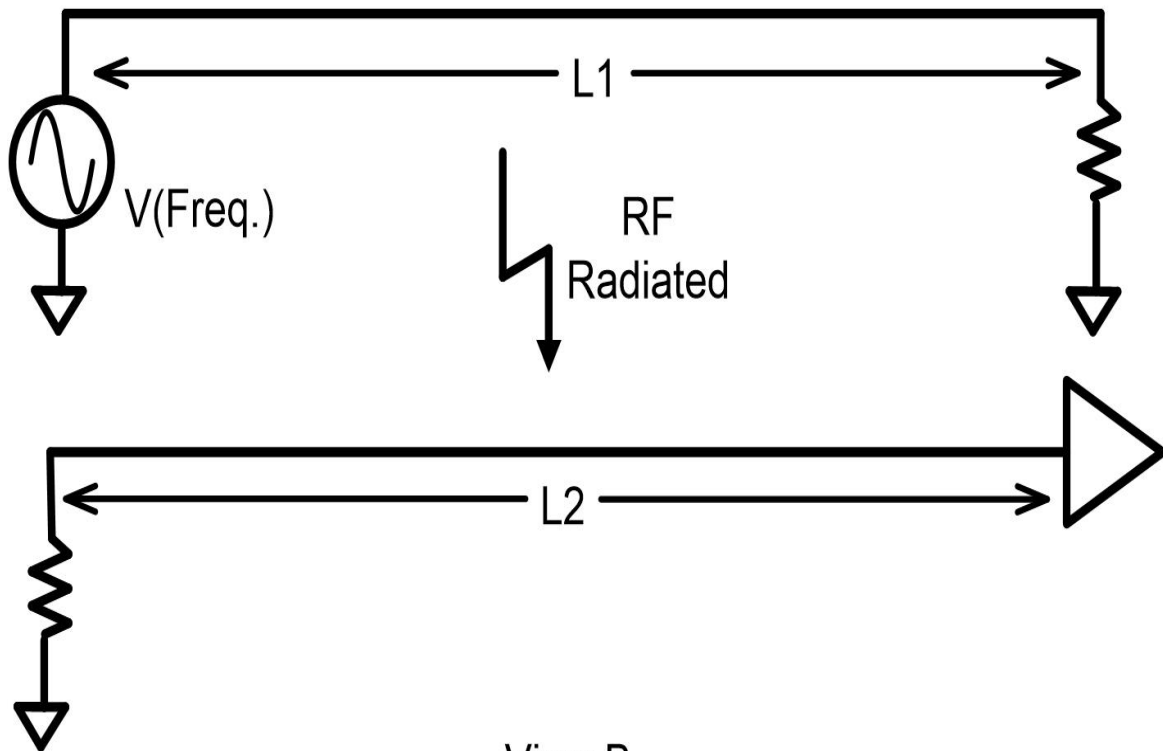
View A above shows a circuit schematic with a digital signal traversing a PCB in close proximity to a sensor connected to an ADC.

View B adds probable signal coupling paths, showing parasitic capacitance (C_{par}) between connections and electromagnetic fields from transient currents traversing A to B, resulting in signal coupling between the two connections.

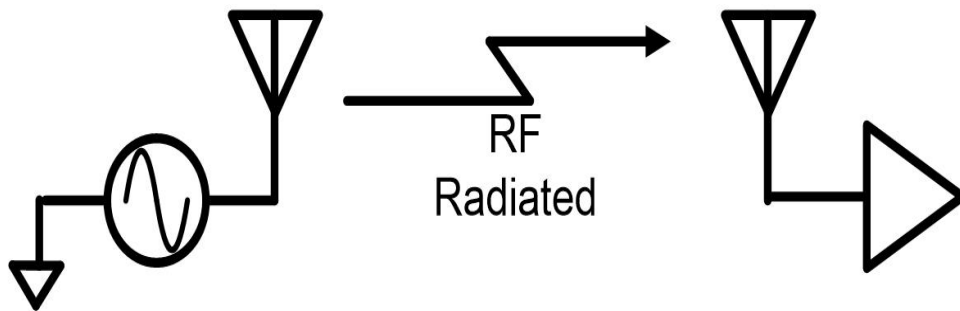
View C completes the story illustrating the physical layout of the connections that are “talking” to each other.

Creating unintentional antennas are another issue (Figure 6-?).

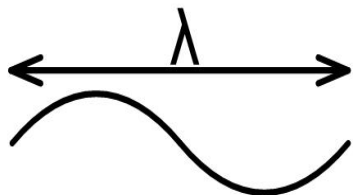
View A



View B



Wavelength = $\frac{\text{Speed}}{\text{Frequency}}$ $\lambda = \frac{C}{F}$



$C \sim 3E10 \text{ cm/s}$

$L < \frac{\lambda}{20}$ For minimal radiation

$\frac{\lambda}{4} = \text{Good antenna}$

Figure 6-10. Noise Coupling through Radiated Methods

View A, shows a signal creating a current along a connection with length L1. Depending on the frequency used, L1 can function as a transmitting antenna, and L2 can function as a receiving antenna. The connection length vs. the signal frequency is important.

View B, shows the unintentional transmitter-receiver pair. The wavelength is a function of the frequency and the velocity of the medium. A good quality antenna is one quarter of the wavelength ($\lambda/4$) and connections shorter than $\lambda/20$ make a poor antenna. Obviously the typical electrical device can create many possible antennas.

Used here, a first order estimate for electromagnetic wave velocity is the speed of light. Although approximate, it is good enough to determine possible noise radiating antennas. As an example:

A 100 MHz square wave clock, will include 7th harmonic content. That 700 MHz sinusoid signal has a wavelength of 43 cm, with 11 cm wires making excellent antennas, and connections longer than 2 cm radiating EMI.

Radiated EMI is especially problematic outside the system, where a system creates noise seen by other devices. Power cords turn into antennas, when common mode noise is on all the power cord wires.

The capability to minimize transient currents in a single conductor can help with radiated emissions. Currents always need a return path however (Figure 6-?).

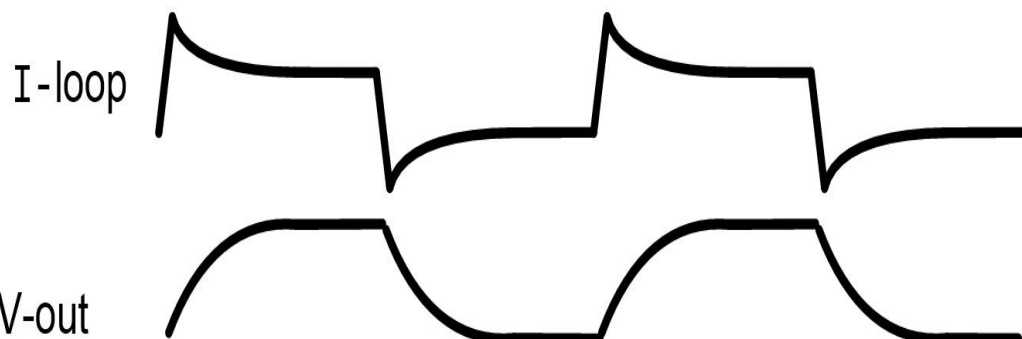
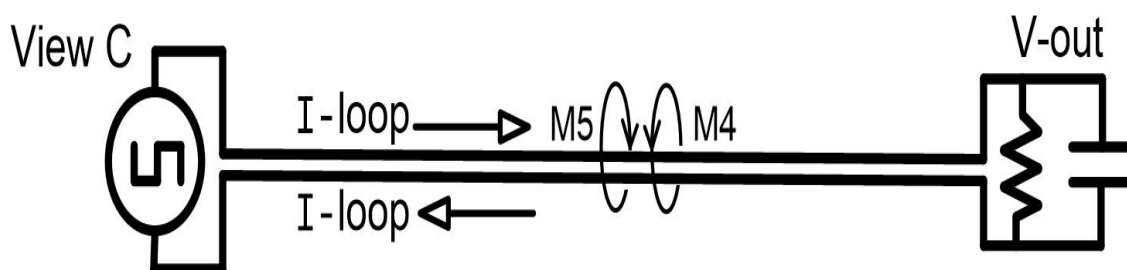
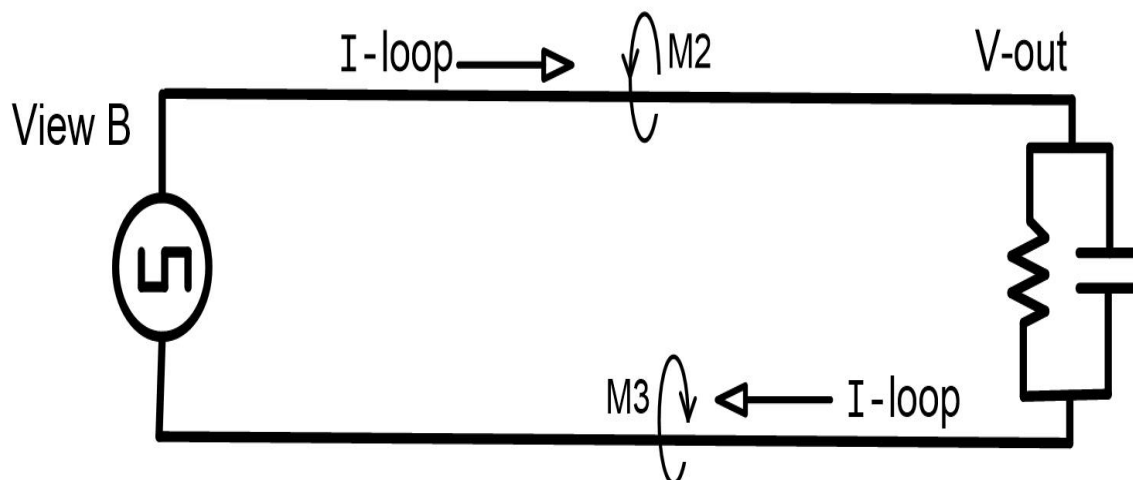
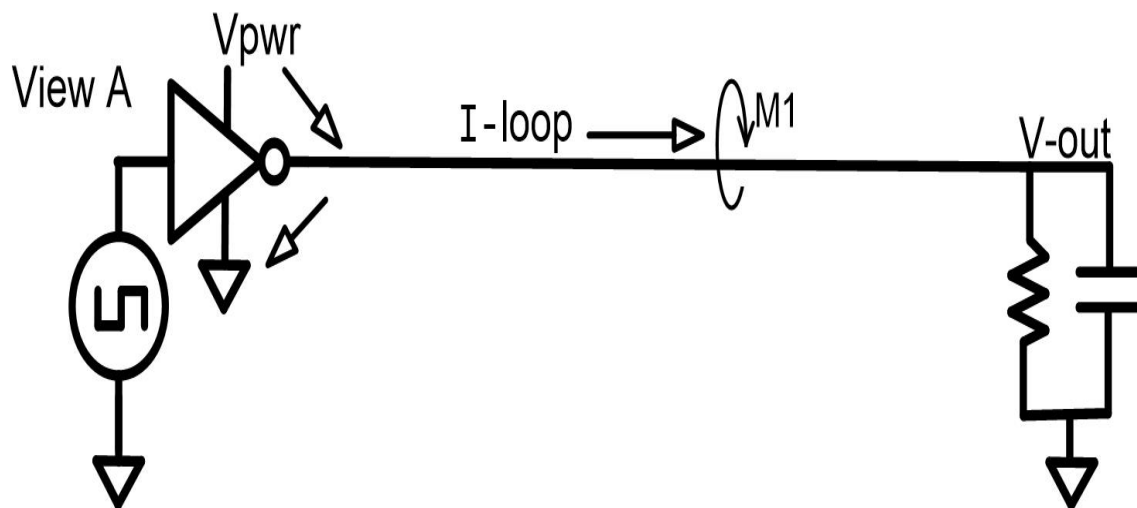


Figure 6-11. Current Loops and Return Current Paths

View A shows the magnetic field (M1) created by a transient current from a digital signal. The I-loop current that goes out in the connection has to return somehow and does so by the ground, creating a current loop.

View B explicitly illustrates the loop's return current. The magnetic fields, M2 and M3 are spaced apart, showing the current loop antenna which has been created.

View C illustrates that closely placed out/return currents causes magnetic fields that cancel each other out. At a distance, the EMI radiated is much less than that radiated in View B. Closely spaced complementary currents are widely used in power and cabling strategies discussed here.

Generally, EMI problems increase with frequency. Some examples readily illustrate the reasons (Figure 6-?).

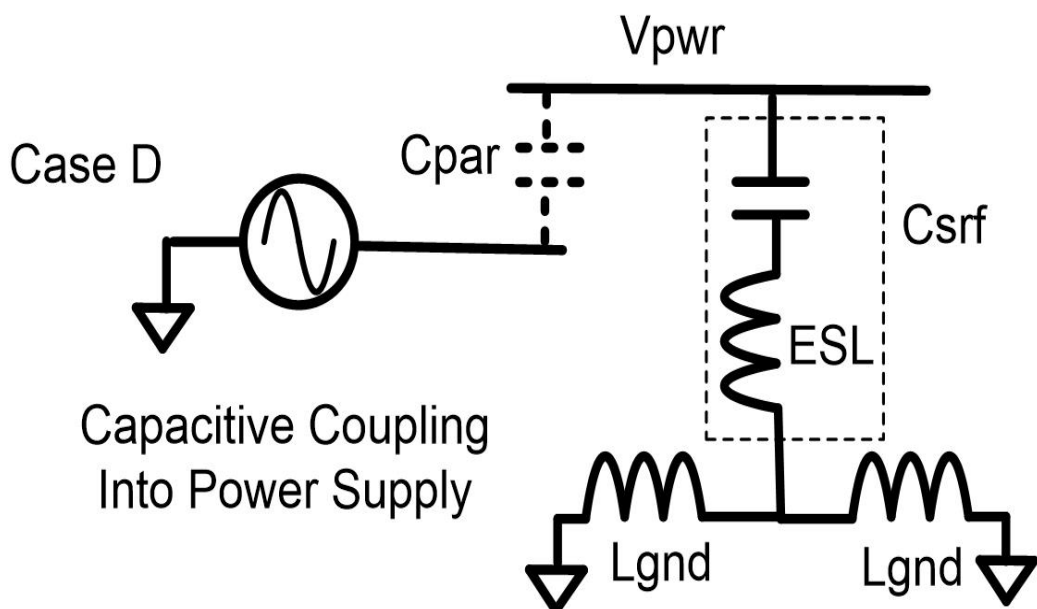
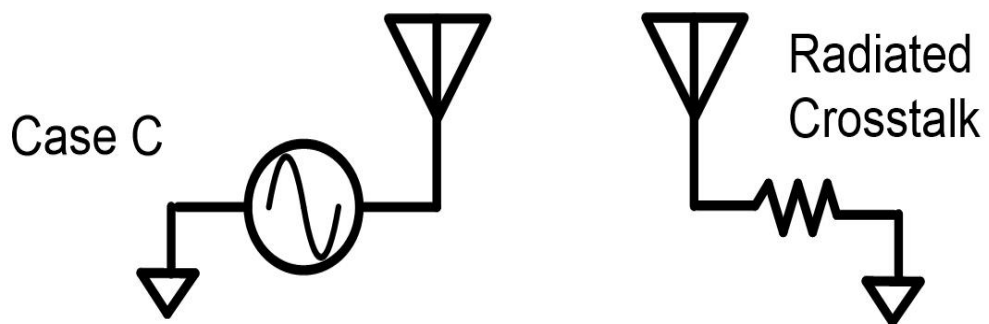
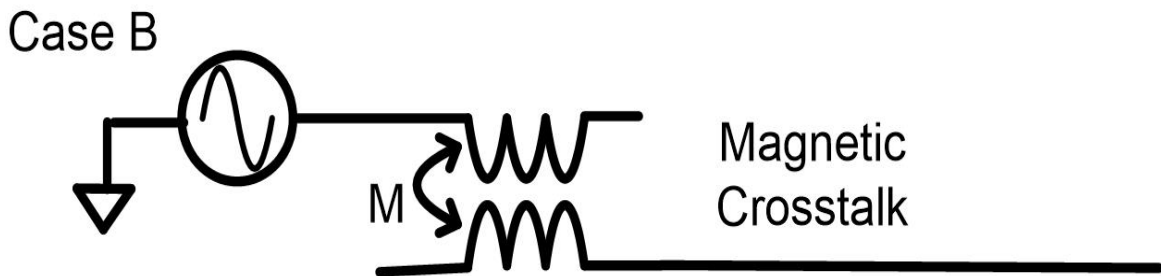
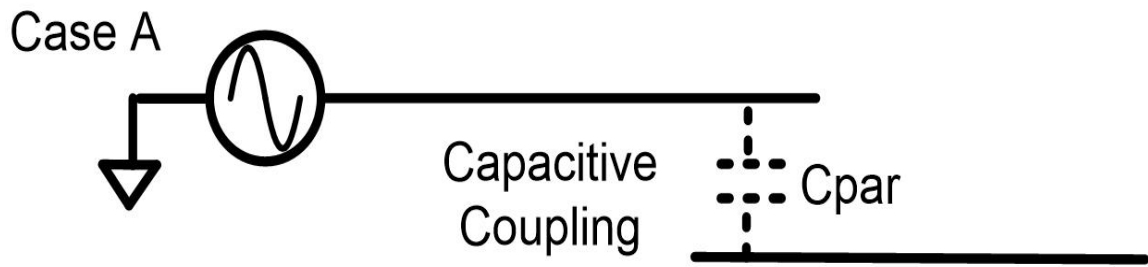


Figure 6-12. EMI Coupling as a Function of Frequency

Case A - Capacitive coupling increases with frequency, due to the impedance of any parasitic capacitance reducing, and signal cross coupling increasing.

Case B - Magnetic crosstalk also increases with frequency. This is due to Faraday's Law of Induction, where the induced voltage increases with frequency.

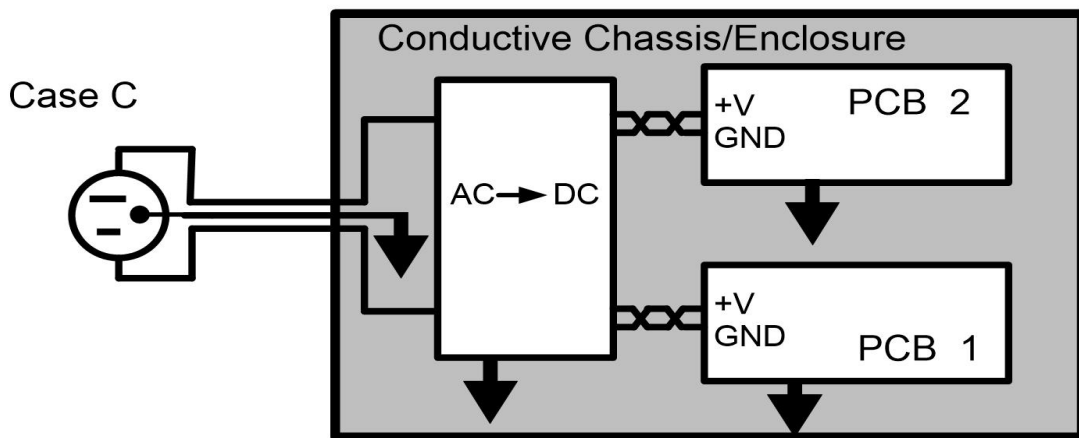
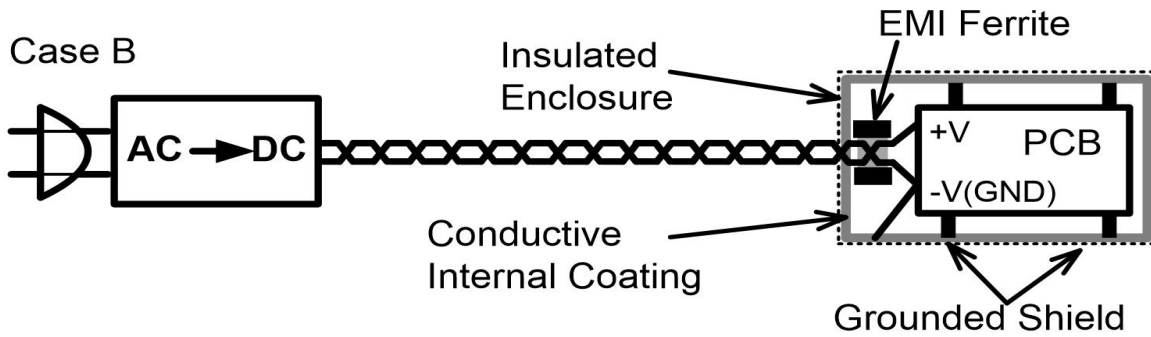
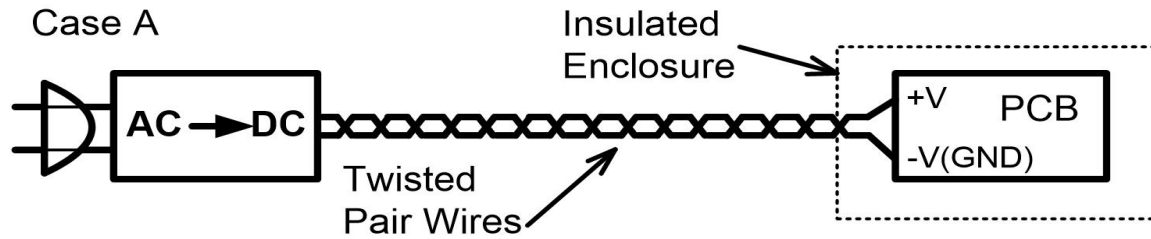
Case C - Radiated crosstalk increases with higher frequencies. Higher frequencies have shorter wavelengths, and more conductors become suitable antennas.

Case D - Power supply filtering becomes less effective at higher frequencies due to non ideal capacitors and the connection inductance. Power filter capacitors (C_{srf}) are not ideal, and exhibit self resonance characteristics due to internal inductance (ESL). Above the self resonance frequency they no longer function as a capacitor.

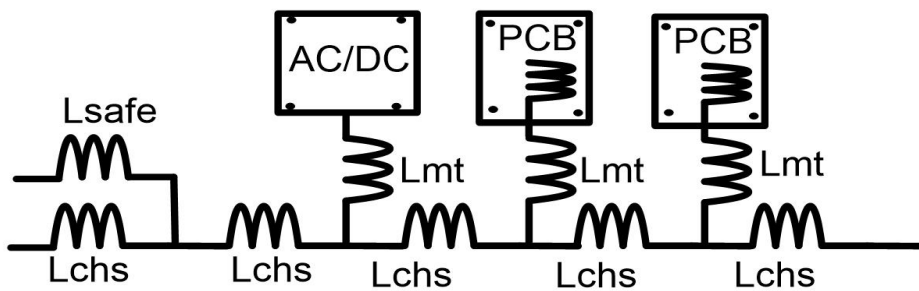
Grounding

A good low impedance grounding strategy is important. Any ground has some impedance and can vary dynamically due to current surges through that impedance network. Poor grounding can affect performance and become more problematic with more complex systems, higher frequencies, large transient currents, or the stresses of ESD events. Establishing a low impedance ground becomes a foundation to close current loops and serves as a stable current return path. A low impedance ground also becomes the common reference point for EMI reduction.

Depending on system size, there are several common approaches (Figure 6-?).



Case C (Ground Connections)



Inductance:

Lsafe - Safety Ground Lchs – Chassis Lmt – Mounts

Figure 6-13. Grounds Within a System

Case A shows a common scenario for small low power devices. The system ground is the PCB ground plane. The AC-DC supply is a wall plug in, and the connection between uses a low EMI twisted wire pair to avoid an open current loop radiator. To be viable, the PCB needs to be low noise, due to no enclosure shielding.

Case B shows another popular approach where the internal PCB generates more EMI and needs shielding. The internal PCB still provides the system ground. A conductive shield encloses the PCB, connected to the PCB system ground, and serves as a Faraday cage to reduce emissions. A ferrite bead around the power wires reduces the common mode noise antenna characteristics of the power wires, while transient current EMI is minimized using a twisted pair. This example is typical of devices like laptop computers and tablets.

A suggested strategy: Creating a design without shielding can be done, that easily converts to the shielded version. If regulatory testing fails radiated emissions, the shielded version is ready to go as a substitute. Conductive coatings for the inside of plastic enclosures are readily available for this purpose.

Case C illustrates the most common grounding strategy used, generally used for larger electronic systems. All devices exist within a conductive chassis and the chassis serves as both common ground and Faraday cage. Due to multiple connection impedances there can be some dynamic variance between all the connected together grounds.

Case C (Ground Connections) illustrates the multiple inductances in the ground paths of Case C. The AC safety ground (L_{safe}) connects to a chassis structure that can have small amounts of chassis inductance (L_{chs}) in it. Ground connections in the mounting (L_{mt}) of each PCB can also have some effect. Multiple grounding mounts for each PCB to the chassis are suggested as a best low impedance practice. A solid conductive sheet ground underneath the circuit boards is greatly preferred. Air vents and external connections can be on the other surfaces of the box. A chassis

ground configured in this manner is generally “good enough” to serve as a solid reference foundation.

Even with a good quality grounding strategy, some dynamic variance will exist across a ground grid. Seeing 100 mV transients from corner to corner of a metal chassis is not uncommon. Poorly designed devices can have much worse. Consequently, some variance in ground needs to be accounted for (Figure 6-?).

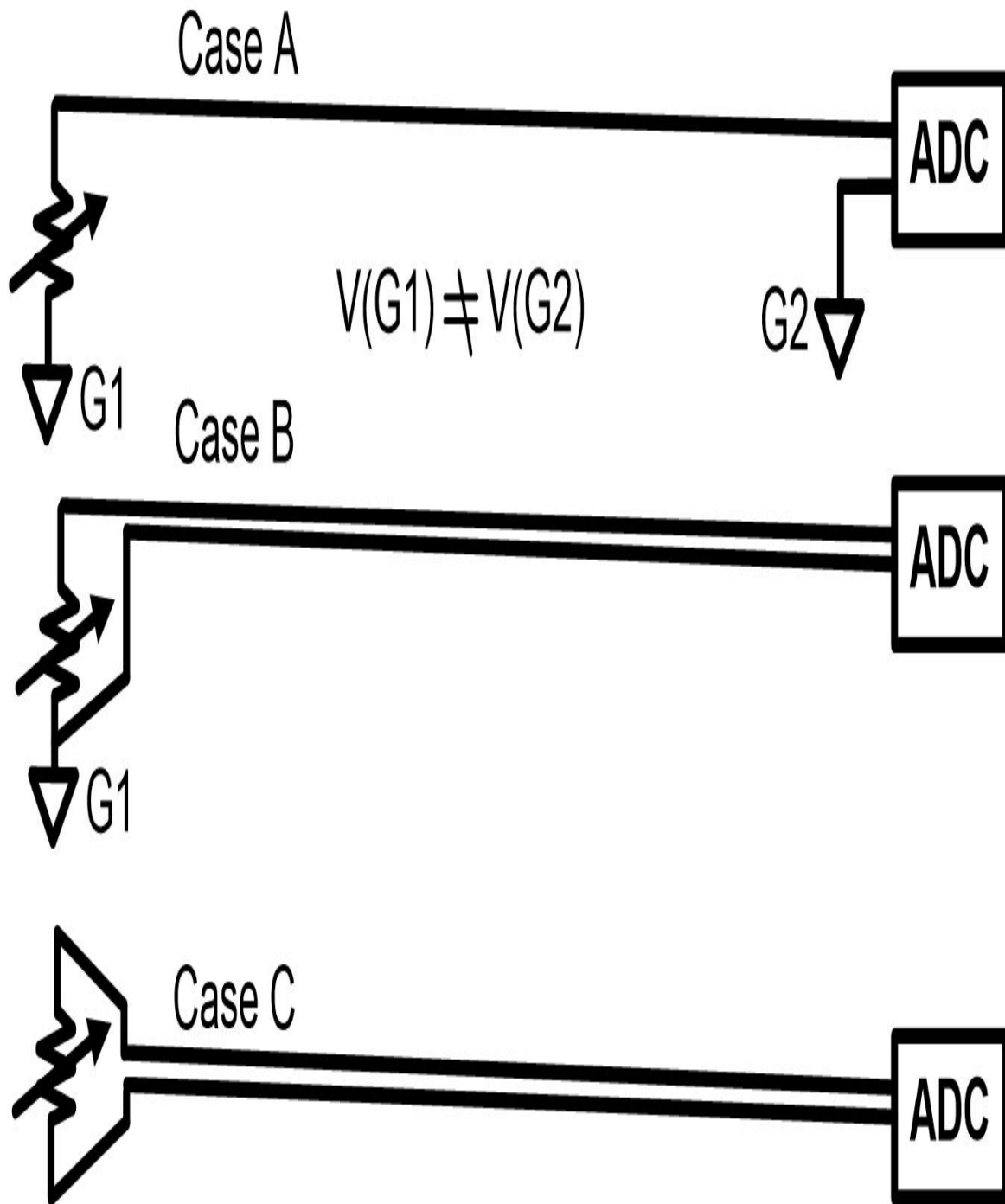


Figure 6-14. Signal Grounds Near Sense Point

Case A shows a sensor, connected to a differential ADC input. Due to ground noise, the voltage at G1 is not the same as G2, and measurement errors will occur.

Case B uses the same ADC, and illustrates the idea of sensing a remote ground in order to remove ground noise errors.

Case C illustrates a sensor that does not need to be grounded. This offers the best option to avoid ground noise and should offer the lowest noise result. As with all analog signals, keeping the connection short is always important.

After creating a low impedance ground at the chassis level, a strategy to do so within each PCB is also necessary:

PCB 4 Layer Cross Section

Top Signal
Insulator
Ground Plane
Insulator
Power Plane
Insulator
Bottom Signal

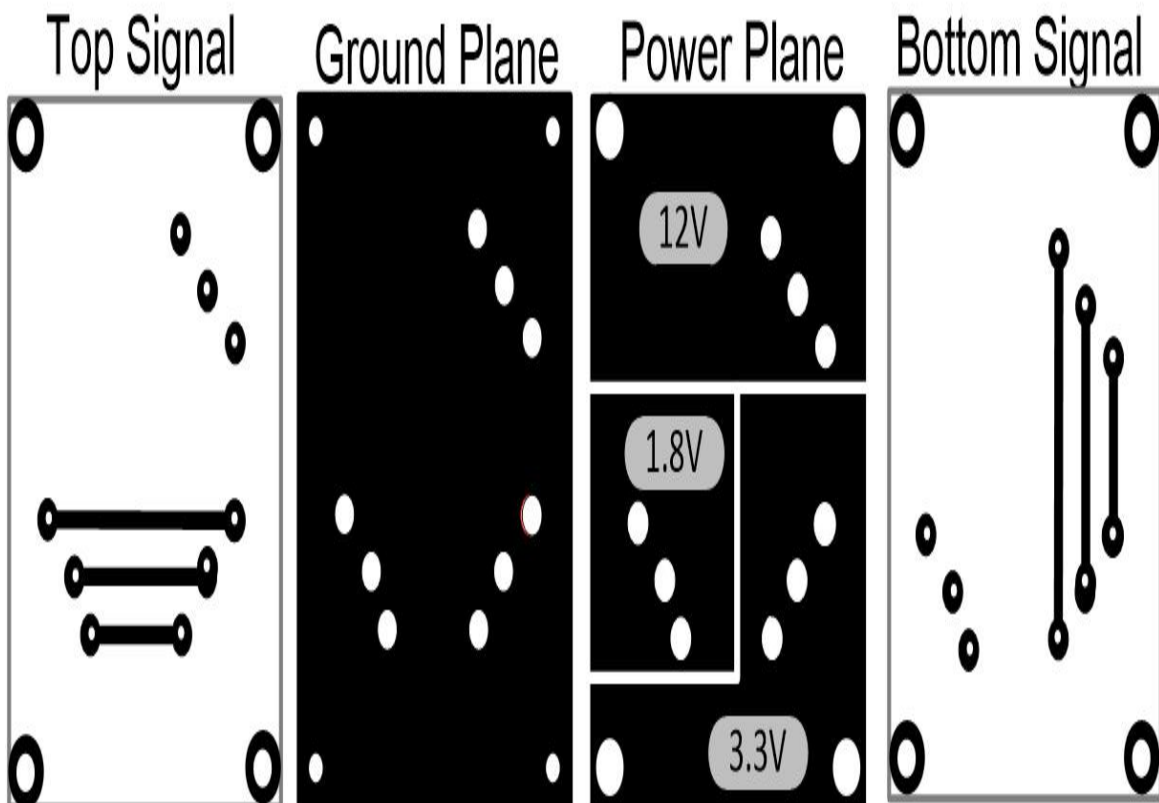


Figure 6-15. PCB Ground and Power Layers

Modern PCB designs use a solid ground layer within the PCB. The most commonly employed method is four layers, where the two outer layers are used for signal interconnect, one internal layer for ground, and one internal layer for power supply distribution. More complex boards may add additional layers.

Earlier generations of PCBs did not use a dedicated ground plane and invariably suffered from reliability and ground noise problems.

Splitting of grounds is not suggested for general use. The only exception is when a portion of the electronics needs to be fully isolated. This comes into use in medical devices, where patient attached electronics must be unattached from any DC current path.

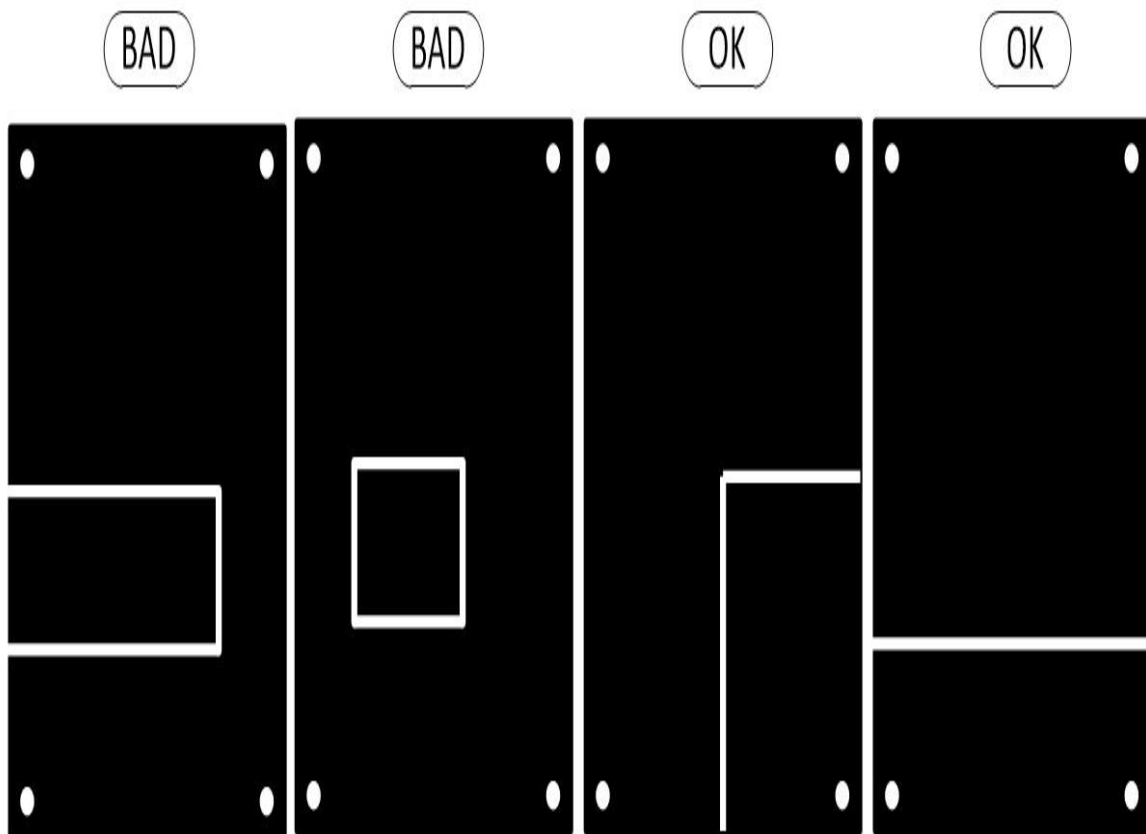


Figure 6-16. Ground Splitting

If grounds need to be split, maintain each of the ground planes as a solid entity, as much as possible. Currents should traverse the ground with

minimal impedance.

Isolating grounds require special support circuitry (Figure 6-?).

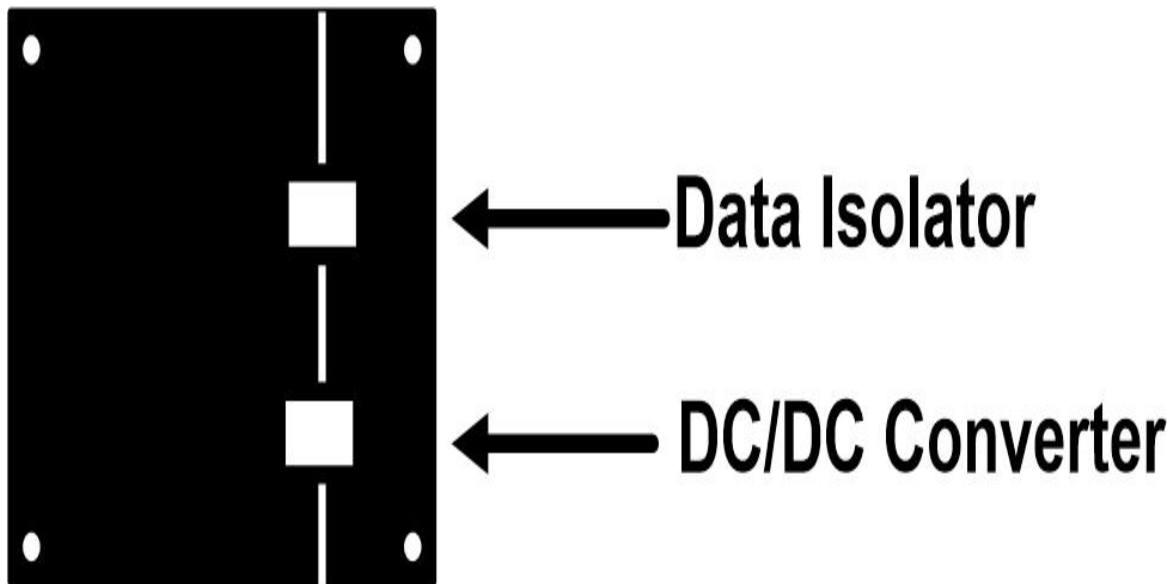


Figure 6-17. Isolated Grounds Typical Structure

For isolated grounds, power needs to be transferred without a DC return path. This is commonly done with a DC-DC converter. Power is transferred by magnetic methods. Data communication also needs to be done with no DC return path, commonly done with an opto-isolator, or other data link that does not share a physical connection.

Again, splitting grounds on a single PCB is not suggested for general use.

Reducing Conducted Emissions to AC Power Mains

Most designers don't develop AC-DC power converters, due to many low cost options already available. One part of the AC-DC converter may need attention however, namely conducted emissions placed on the AC mains input. This is largely done by filtering methods, with a typical filter shown in Figure 6-?.

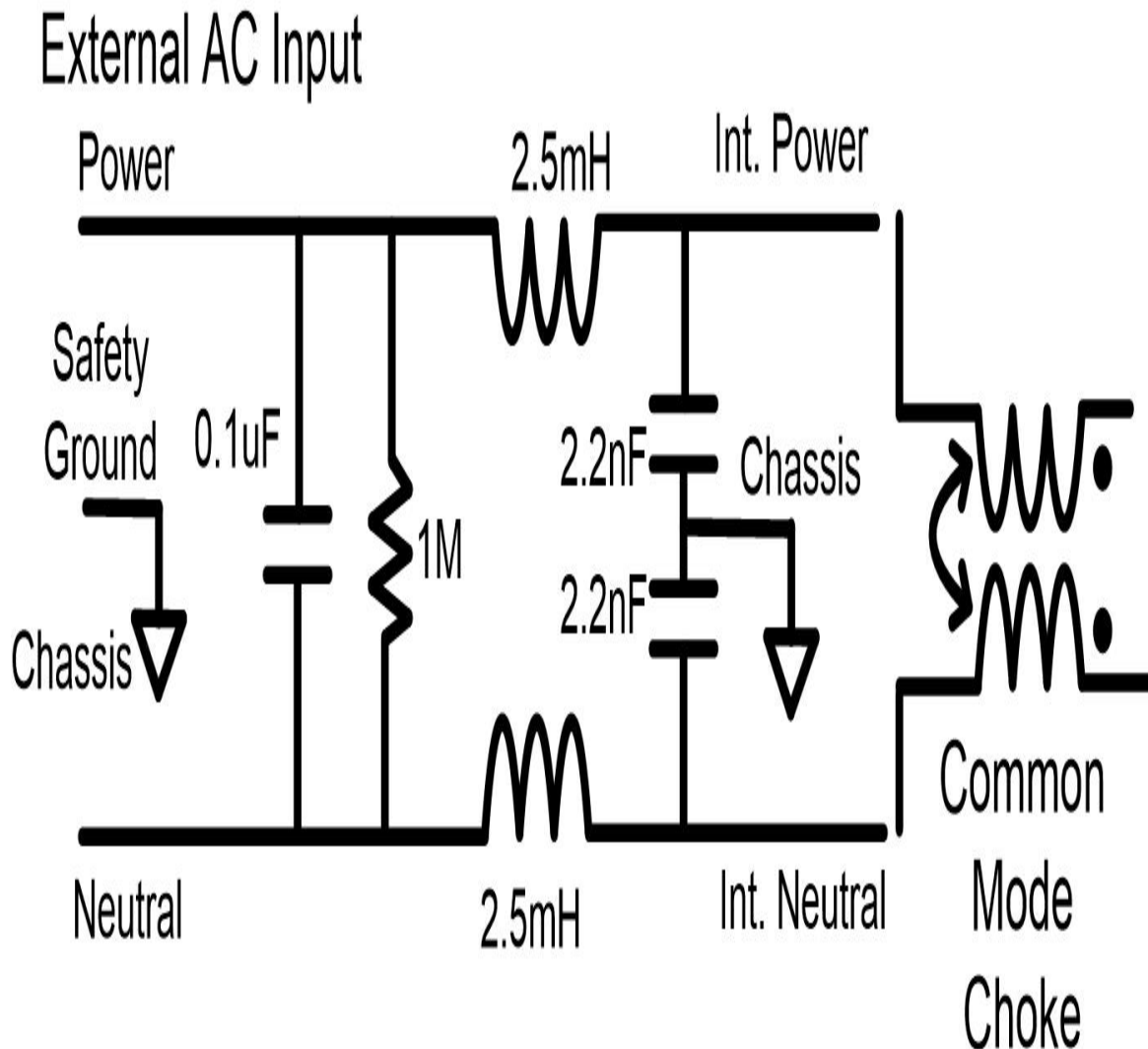


Figure 6-18. AC Connection Suppression-Filter Configuration

This commercially sold filter is typically available within a chassis mount power connector. Some include a common mode choke (right side as shown) and others do not. If conducted emissions become a design issue, a power plug connection can be purchased that includes the filter.

Cable Interconnect Strategies

Cabling interconnects can have different strategies, depending upon the application (Figure 6-?).

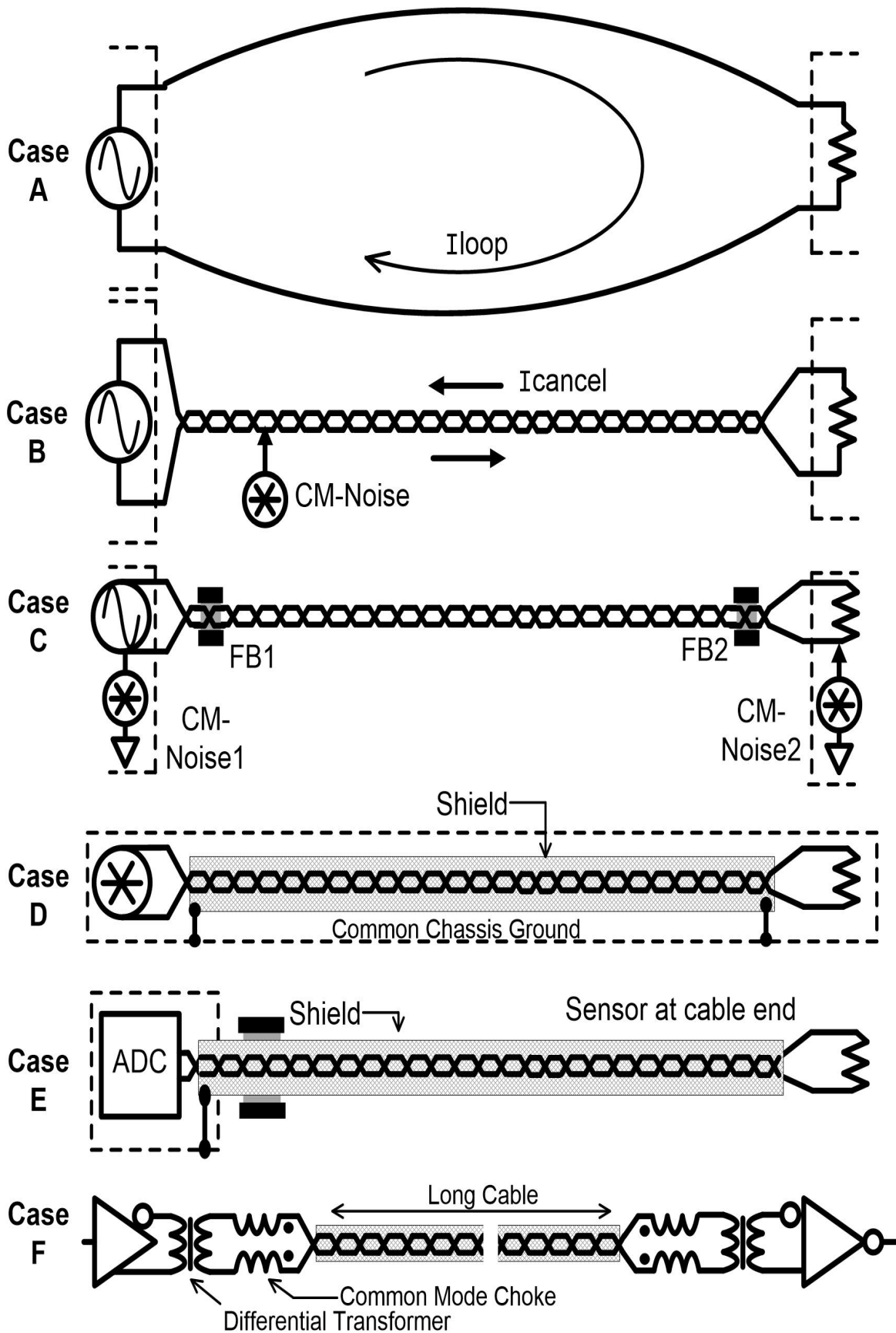


Figure 6-19. Cable Interconnect Strategies

Case A is the open current loop discussed earlier. Open current loops should be avoided due to creating radiated EMI, and the lack of balanced and equal paths on both sides of the signal.

Case B introduces twisted pair wires that minimize radiated EMI through magnetic field cancellation. With closely spaced, equal and opposite currents, the net field generated is minimized. Additionally, most EMI coupled into the twisted pair will be common to both wires, creating “common mode” noise. This allows a differential receiver to improve EMI rejection by subtracting the signals from each other. For power connections, and differential signals, this approach works well.

Case C deals with common mode noise introduced to a cable. This can be an issue at both the sending (CM-Noise 1) and receiving (CM-Noise 2) ends. Noise introduced to all wires in a cable bundle creates an antenna to radiate noise. The use of a FB at each end effectively separates the noise source from the antenna. This is the preferred method for any cable connection going outside a shielded enclosure.

Case D looks at a special case where a cable inside a system is a noisy radiator. Using a braided shield over the conducting wires and connecting both ends of the shield to the common chassis ground will reduce radiated EMI from the cable.

Case E is suggested when connecting to a sensor or other low amplitude signal source on a cable beyond the system enclosure. A shield over the cable, connected to the ground at the receiver (ADC here) minimizes noise seen at the receiver. A ferrite on the cable minimizes the cable acting as an antenna.

Case F is common for LVDS data communication on long cables. DC Isolation is created with transformers at both ends of the cable. This eliminates a need for a shared ground between the devices at each end. Common mode rejection chokes are also included, rejecting the noise

pickup of any lengthy cabling scenario. This cabling structure exists within all Ethernet cable connections.

Twisted pair wires are more important than just holding two wires together. With opposing currents they reduce EMI creation and avoid open current loops. With tightly twisted wires, and differential signals, they reject common mode noise when using a differential receiver.

The cable strategy for a sensor that requires both power and signal becomes two twisted pairs, one for signal and the other for P&G. Cable structures with multiple twisted pairs are commonly used.

Reduce Noise Generation at the Source

Noise reduction starts at the sources. Selection of circuit methods, clock frequencies, signal distribution, and others techniques will make significant noise reduction in any system.

The predominant EMI sources in most systems are devices that switch current or cause high frequency transients. All digital devices, clock tree drivers, switching mode power supplies and high current switching amplifiers are common EMI sources. As frequency and currents increase, so does the EMI.

The first EMI source easily improved is the transition time of digital devices (Figure 6-?).

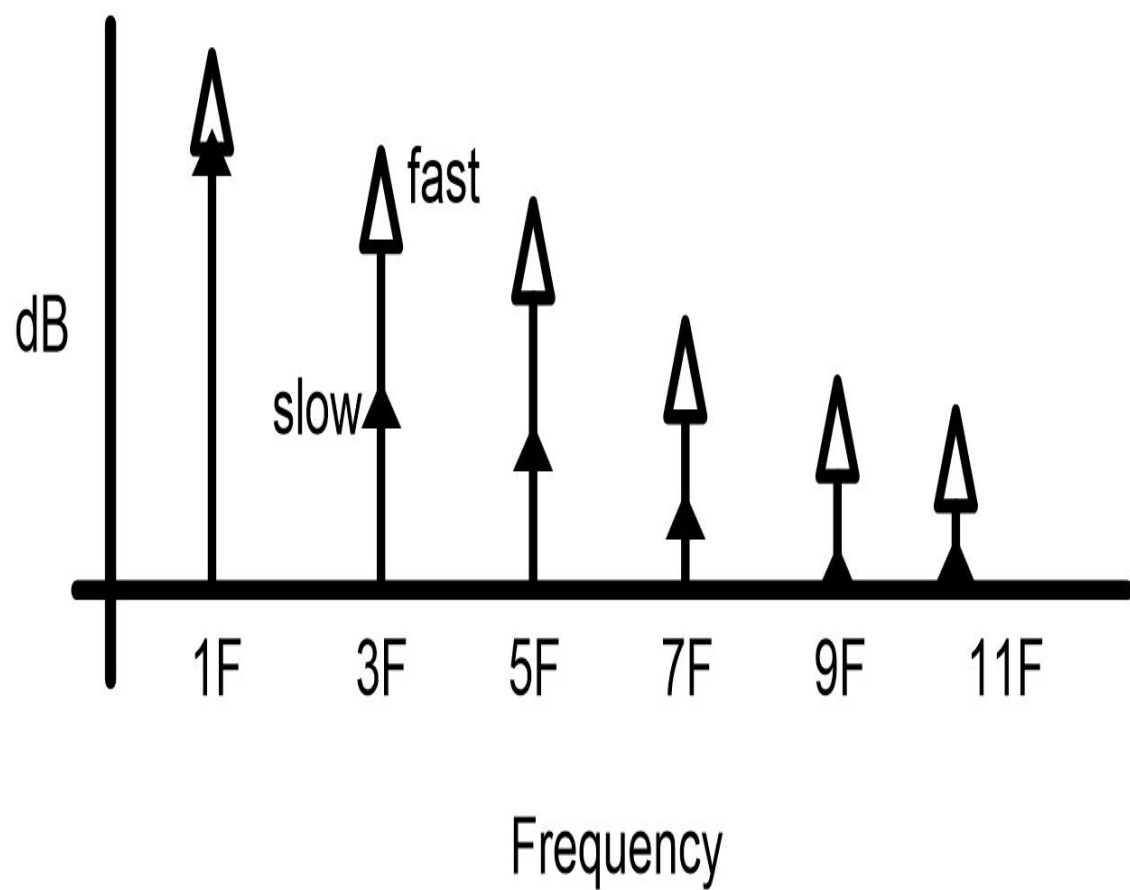
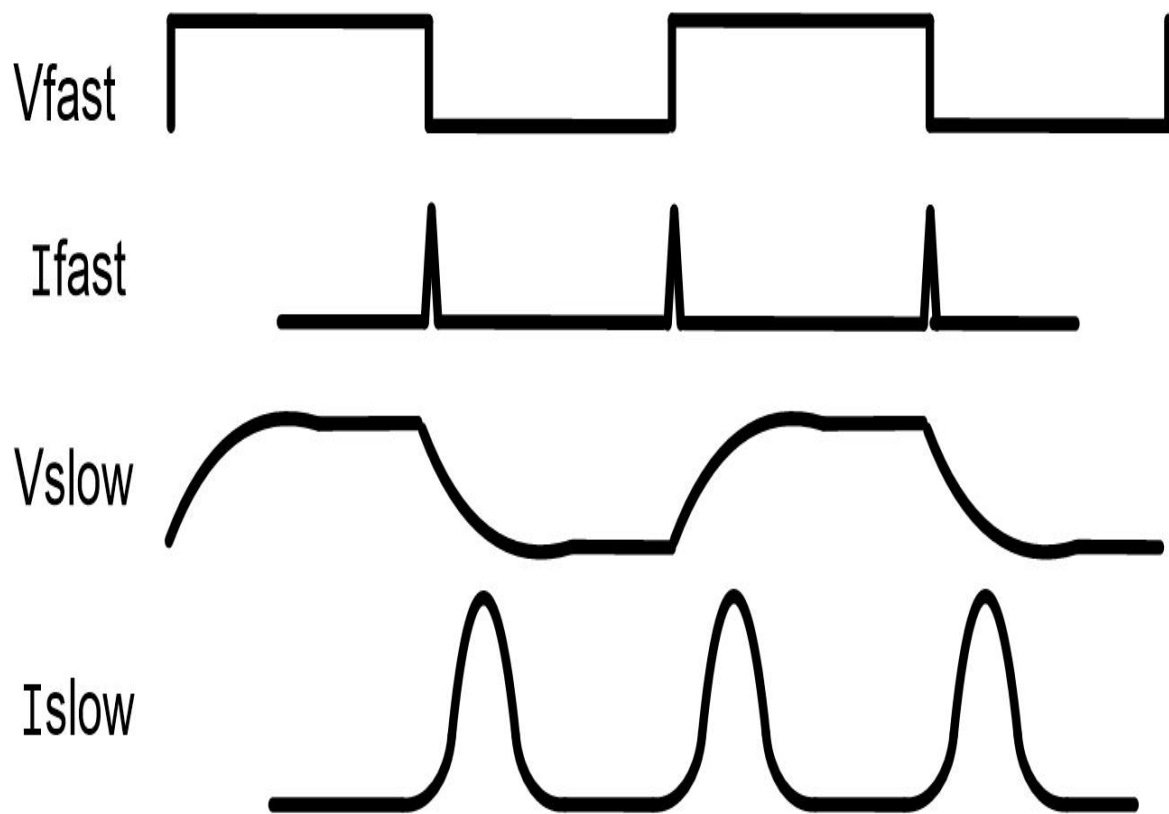


Figure 6-20. Digital Rise time vs. EMI

Fast transition edges, and the resulting current spikes, produce more EMI than slower transitions. With slower rise/fall times, harmonic amplitudes are reduced. Changing the I/O driver strength is the technique available to control transition speed. Digital ASIC's, FPGA's, and CPLDs often have the capability to set driver strength for each output. Some MPU and MCU devices also provide this feature. Look for a configuration option to set "drive strength" when programming the device. Slower data ports can use a weaker drive strength setting that reduces EMI and avoids the need for transmission lines.

Many embedded controller systems are designed using needlessly fast clocks and quick data transients to control slowly executed physical tasks. This is common with mechanical devices under electronic control. Slower clocks and "softer" logic edges result in both lower EMI and lower power consumption.

Using LVDS methods for fast/distant digital data has EMI advantages for several reasons (Figure 6-?).

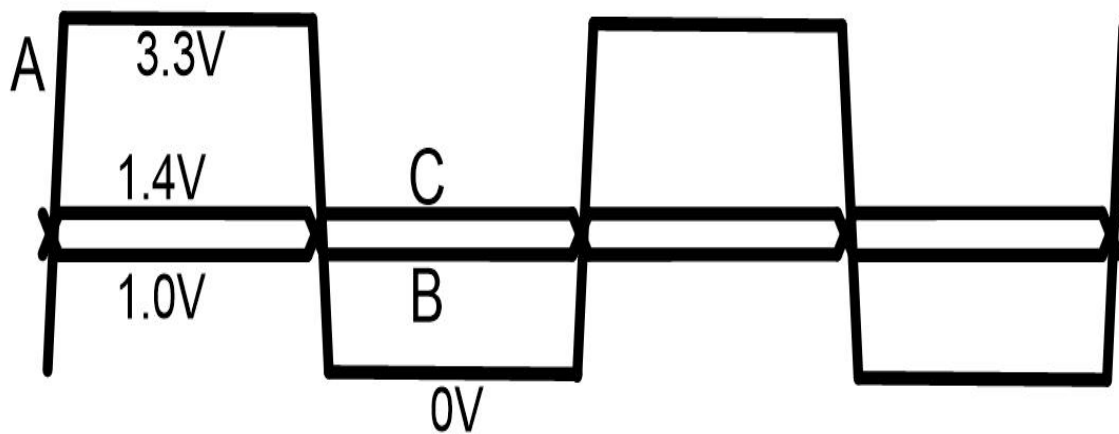
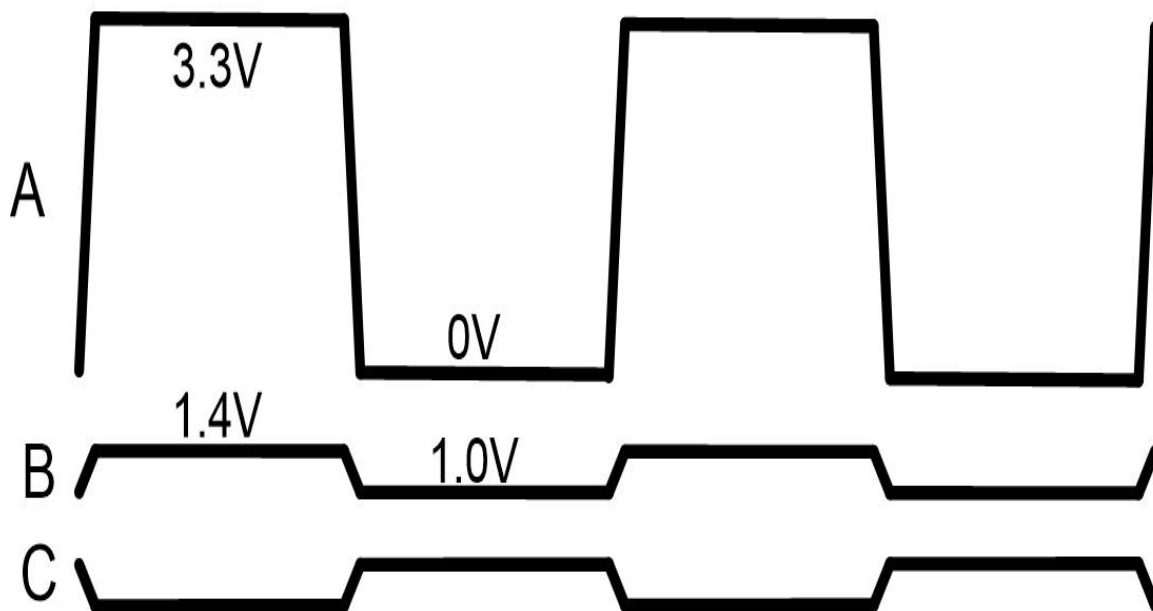
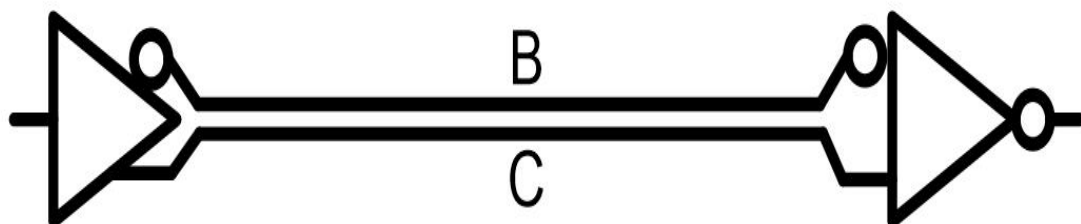
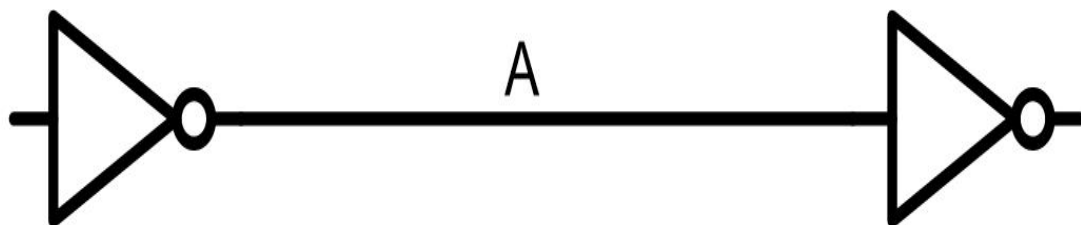


Figure 6-21. LVDS methods vs. EMI

LVDS uses complementary currents, similar to the twisted pair wires described earlier. There are additional advantages, illustrated here. The typical logic and LVDS voltages for 3.3V power devices are shown. Typical CMOS (signal A) uses 3.3V high and 0V low. LVDS (signals B, C) use smaller amplitudes, 1.4V high and 1.0V low. This is possible because differential signals are largely immune to P&G noise, and reject common mode EMI. Consequently smaller amplitudes work reliably and create less noise.

Synchronous digital devices create large amounts of EMI at harmonic multiples of the clock frequency. Reducing clock EMI can be done by spectral spreading (Figure 6-?).

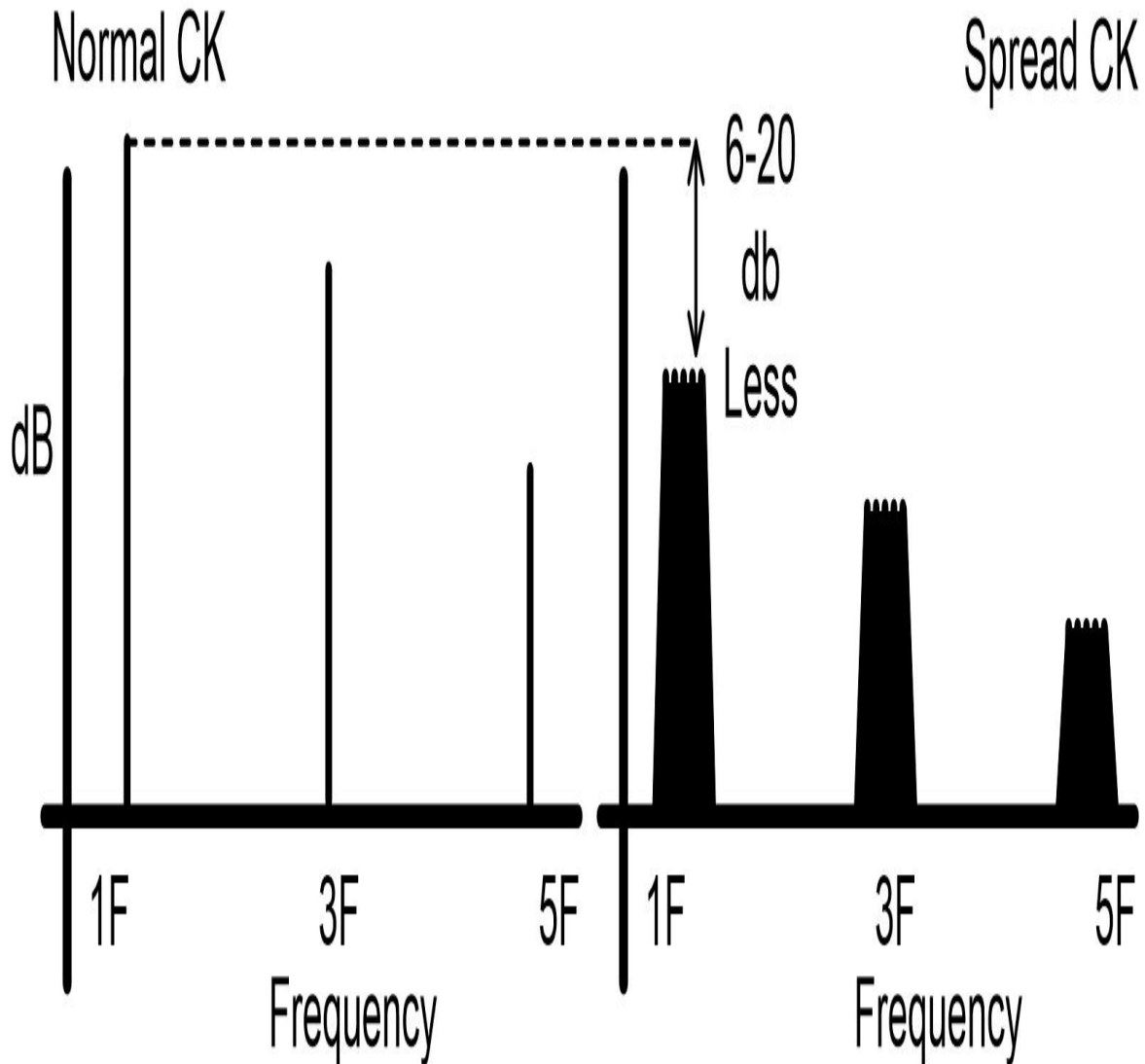


Figure 6-22. Spectral Spreading of Clocks

Also known as “clock dithering” or a “spread spectrum clock generator” this works by changing the clock period slightly on a cycle to cycle basis. The change is small, and does not violate the setup-hold timing of devices. Examining the clock in the frequency domain, shows a significant change. The clock signal was a narrow single frequency spike, and has become a wider and lower amplitude “spread” output. Improvements greater than 20 dB are common. Many devices have failed FCC emissions testing with a conventional clock, and passed with a spread clock.

Any digital device still needs low impedance P&G, aggressive power decoupling, and other EMI and signal integrity issues applied.

Spread spectrum methods are sometimes applied to SMPS to reduce emissions from the switching voltage regulators. But, spread spectrum methods are usually not needed for SMPS. The most effective methods for SMPS noise reduction focuses on snubbers and local isolation (Figure 6-?).

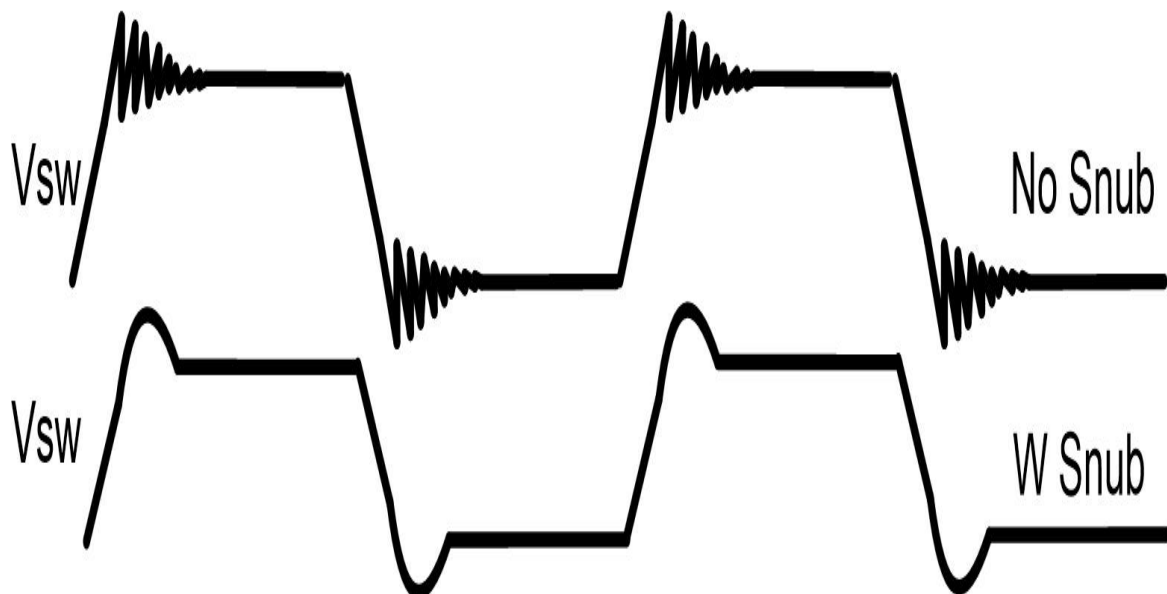
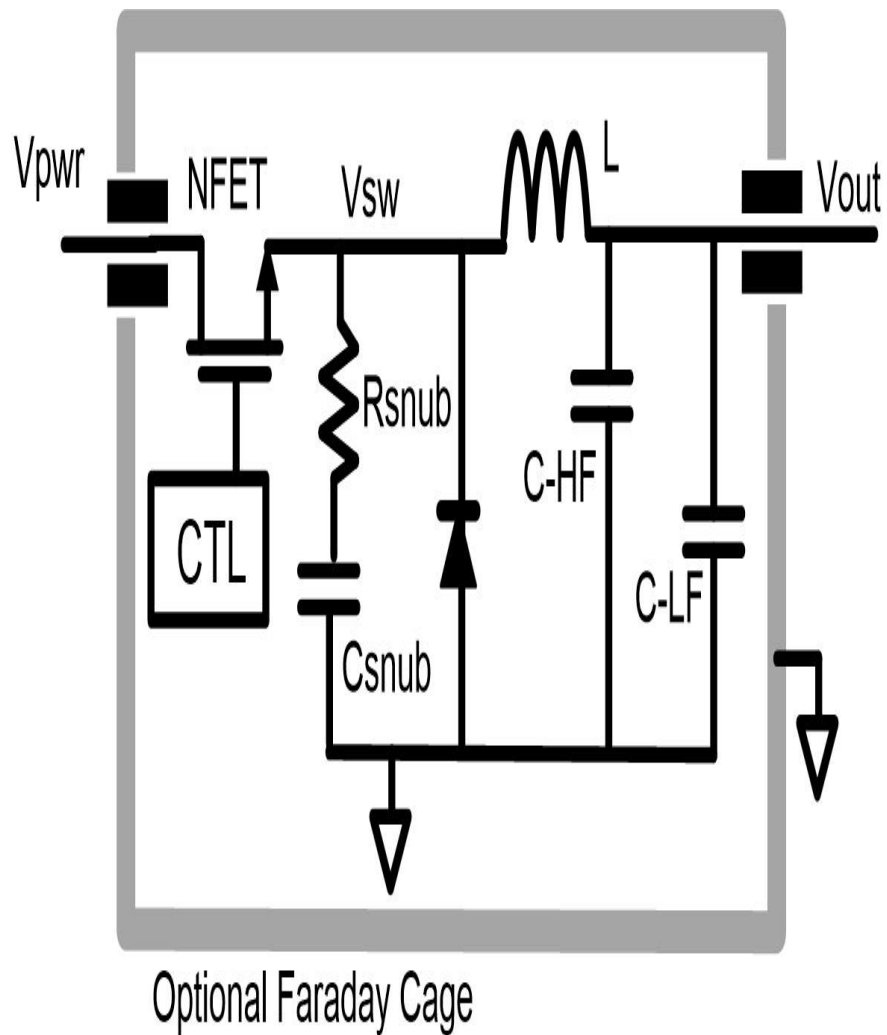


Figure 6-23. Localized EMI Choking in SMPS and RC Snubbers

SMPS should always include PCB locations for a few simple noise abatement techniques. Shown above is one example of a step down SMPS, where the node “Vsw” is the largest noise source. Due to constant on-off switching of the NFET, the Vsw node sees heavy transient currents that create EMI.

Attaching a LPF to the Vsw node, often called a “RC snubber,” reduces EMI significantly by creating a low impedance path to ground for high frequencies. The RC time constant of the snubber is selected such that the value is one tenth of the minimum functional switched pulse time period. This suppresses high frequency transients without changing basic functionality.

As an example, a SMPS uses a controller with a minimum switched period of 1E-6 seconds. That implies a snubber with an RC time constant of 1E-7 seconds. For $R = 10$ ohms, the appropriate $C = 0.01\mu\text{F}$. Generally, vendors of SMPS controller chips can provide application notes on RC component selection suitable to their specific ICs.

In addition to the snubber, several other items are suggested. The output capacitor, (C-LF), is not ideal, and often exhibits poor high frequency characteristics due to large internal inductance. Adding a supplemental parallel device (C-HF), with better HF characteristics will reduce HF emissions.

Placing FBs (or other RF chokes) at the input and output of the SMPS reduces noise that travels out to the rest of the system. Using a Faraday cage around the SMPS may be needed in some cases, especially as the power exceeds 50W.

All SMPS designs should include: RC snubber, C-HF & C-LF, and the FBs. If needed, a Faraday cage, and better RF chokes can be considered.

Some items can become unintentional antennas. A common antenna seen is created by a floating heat sink on a power transistor or a heat sink on a MPU (Figure 6-?).

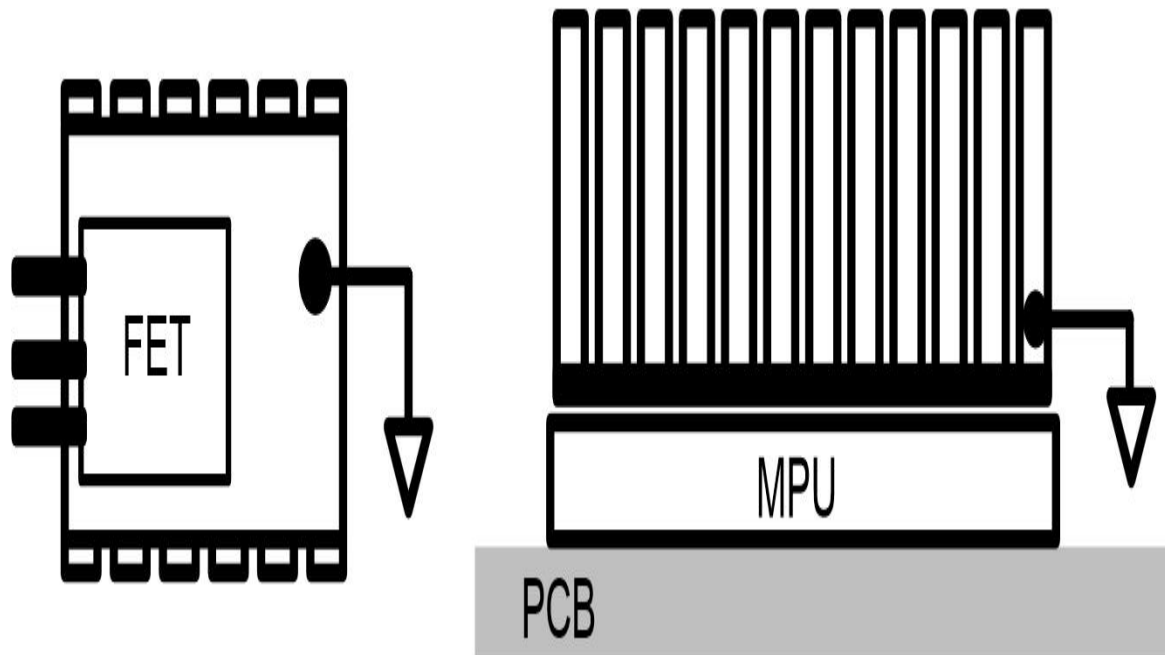


Figure 6-24. Heat Sinks as an EMI Antenna

If the heat sink is not grounded it can create an antenna. Consequently grounding the heat sink needs to be investigated.

Using a filter on motors can also help (Figure 6-?).

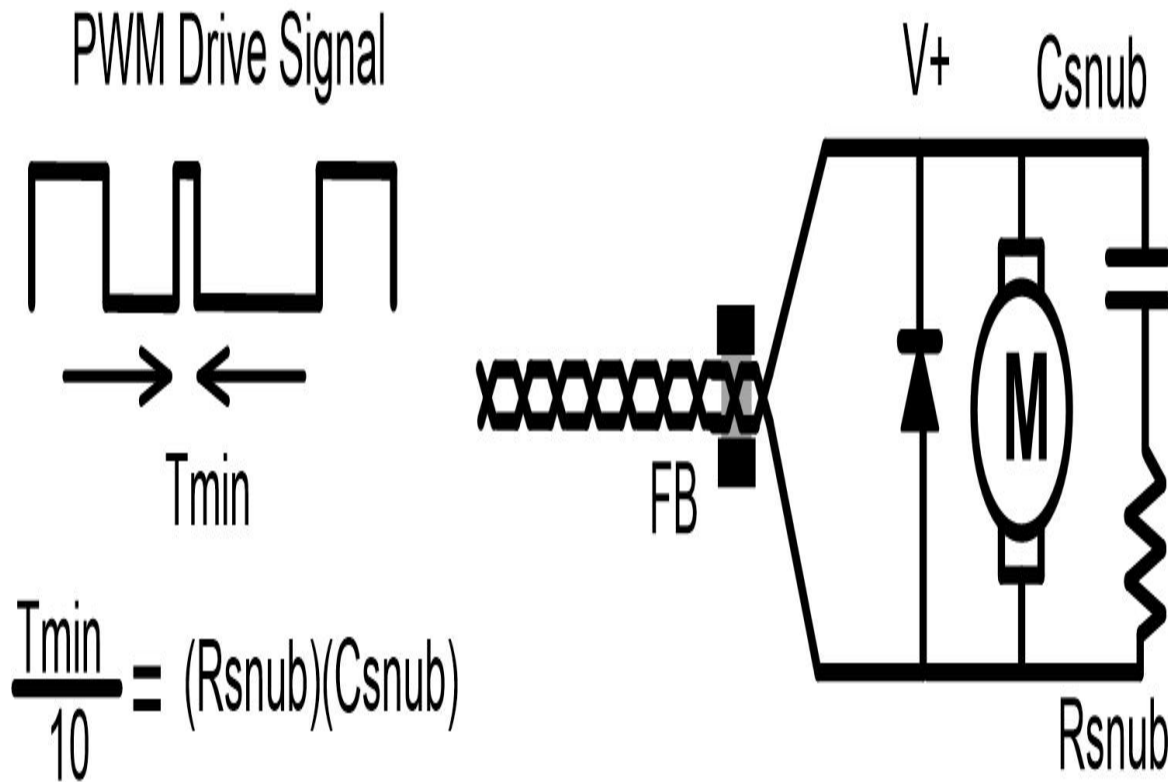


Figure 6-25. Motor Drivers, RC Snubbers for EMI

For a brushed DC motor, using an RC snubber will reduce EMI from both arcing motor brushes, and the voltage transients created by motor inductance. Pulse Width Modulation (PWM) is commonly used to control motor speed and a RC snubber can be selected that does not interfere with PWM functionality. Using FBs at both ends of the twisted pair driver cable also helps to keep the cable from acting as antenna.

Reduce Noise coupling between on board devices

After reducing noise at all sources, the next step is to minimize noise between devices. A quick review of coupling mechanisms between devices is shown in Figure 6-?.

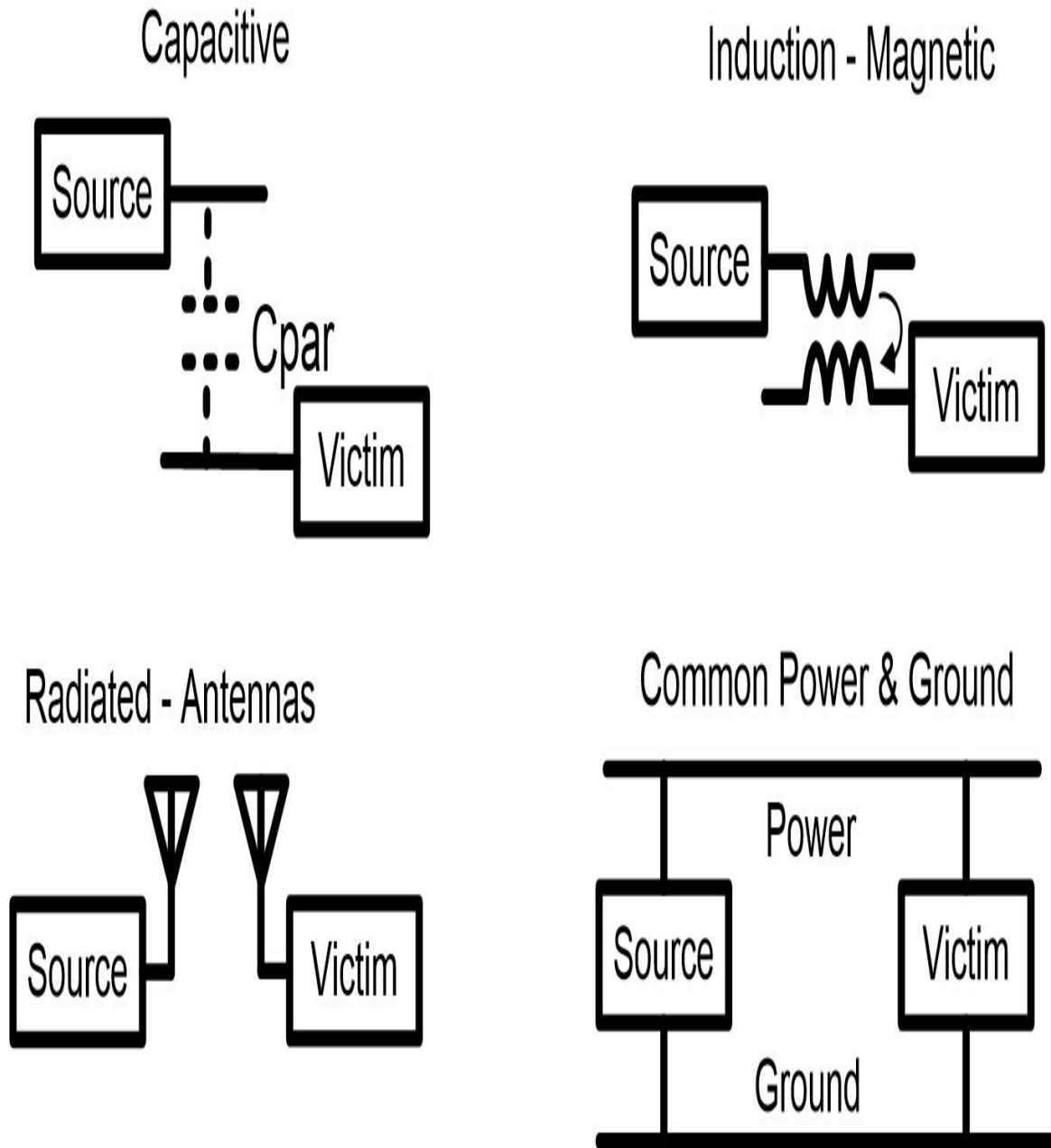


Figure 6-26. Principle Noise Transmission Methods

Digital signals can tolerate limited amounts of noise and priority should be given to protecting analog signals in the system. Incorporating that idea is a good starting point strategy.

Capacitive coupling is often the most common path of noise on a PCB. Noise transmission by parasitic capacitance can be reduced by selective

spacing of sensitive connections, placing grounds between signals, and managing connection impedance.

Magnetic induction noise can be avoided by using digital signals over distance, differential signals, and avoiding long/close parallel paths to signals that have high current transients.

Radiated noise is less problematic within a PCB due to fewer suitable antennas vs. the long wires associated with going off the PCB. Radiated noise from any PCB is easily contained with a Faraday cage.

P&G connections between components provide an easy noise path, thus the need for low impedance ground plane under all circuits and aggressive use of power supply decoupling throughout a PCB

On a PCB, capacitive coupling and P&G noise tend to be the most common problem.

Floor planning, component locations and signal path routes all strongly affect noise performance. To properly deal with this, first understand what components and signals are noisemakers (sources), and, what components and signals are sensitive (victims). Also, determine what signals need to travel long paths on the PCB and which signals are noisy with lots of transition switching. Digital control lines that have minimal activity can be recognized as a quieter class than high speed data pipelines.

Classifying connections from the most sensitive (victim) to the loudest (source) would look like:

- Small amplitude, analog, ground referenced signals (very sensitive)
- Small amplitude, analog, differential signals
- Other analog signals
- Low Activity (static) digital control lines
- LVDS digital data streams
- Ground referenced digital data streams

- Ground referenced data streams going off the PCB (most noise)

The general rule is to keep sensitive items (top of the list), away from loud noise generators (bottom of the list).

In a similar manner, the noisiest chips are switching devices namely digital chips and SMPS devices. The higher the clocking frequency, generally the more of a noise problem it presents.

Power supply connections are not part of this. Generally, power is on its own connection layer and SMPS circuitry is tightly grouped together avoiding lengthy connections. Keeping the analog and digital power supplies separated is suggested. Digital devices create large amounts of transient noise directly on the P&G (Figure 6-?).

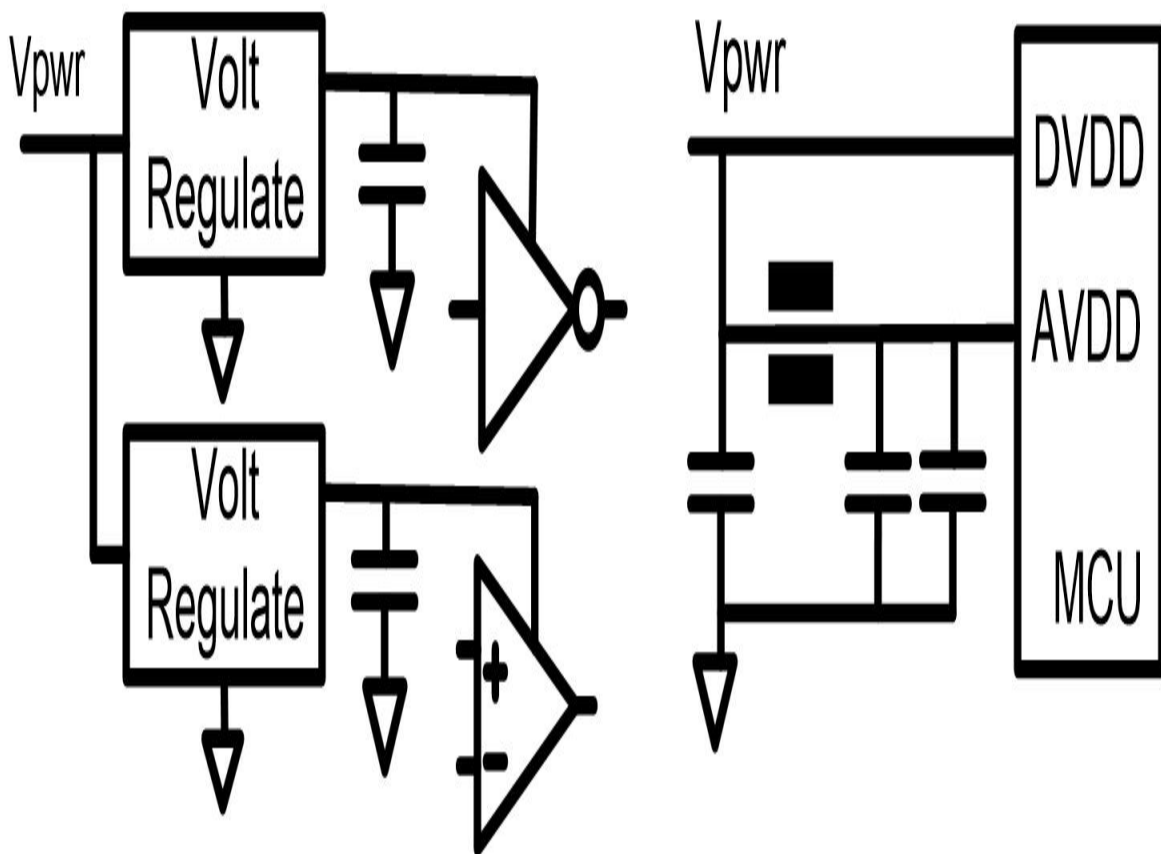


Figure 6-27. Independent Power for Analog-Digital

Analog, mixed signal and any AFE circuits should not share a power supply with digital devices. Using separated power supplies, created with

independent regulators and filters, is the preferred approach. At a minimum, separating analog and digital power (AVDD, DVDD) with a LPF is suggested. Some MCU vendors suggest a ferrite bead and capacitors to keep AVDD quiet, but it's often not enough. Analog circuits are sensitive to power supply noise and perform best with independently regulated and filtered power.

With a good understanding of what's noisy and what's sensitive, a strategy for floor planning, component placement and routing can be made. The first priority in this should be noise (Figure 6-?).).

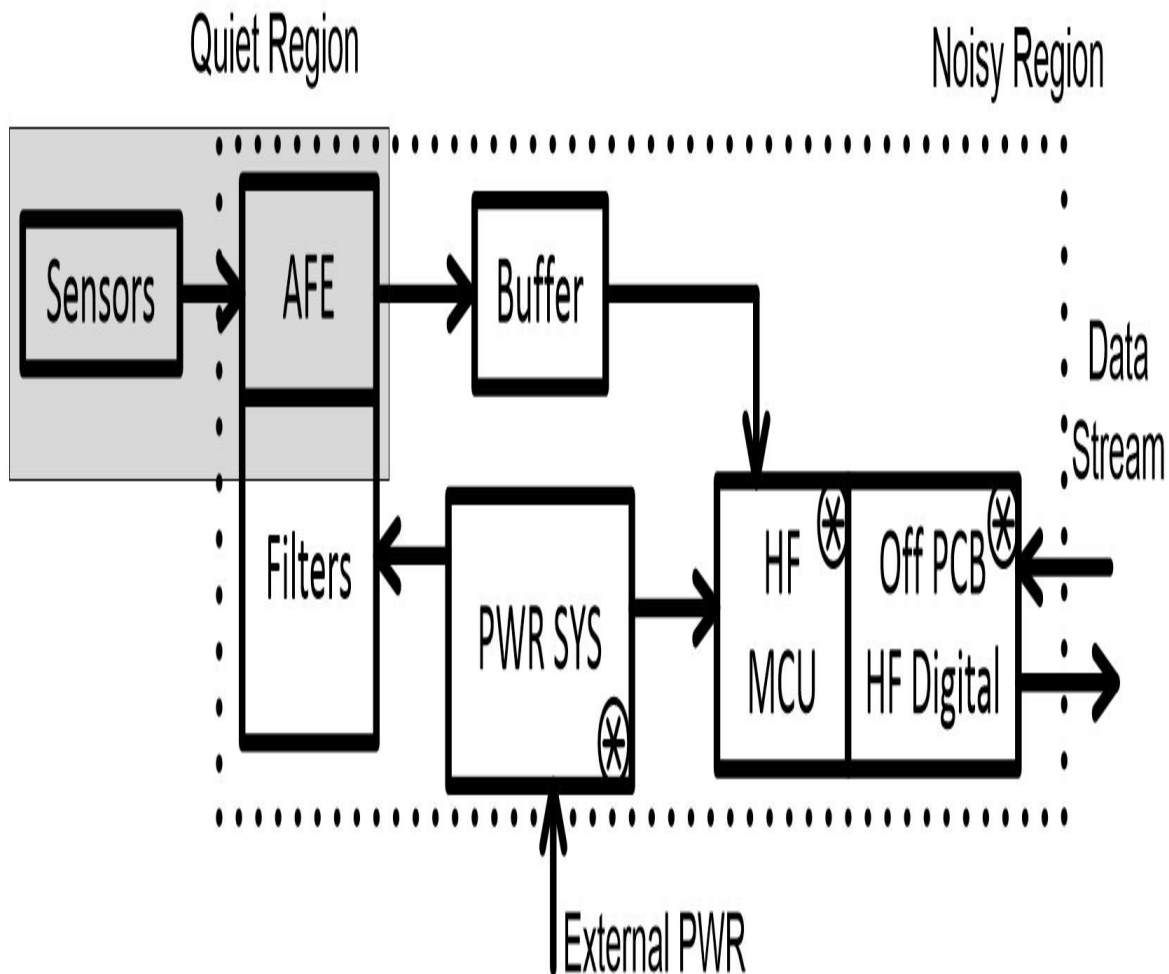


Figure 6-28. Floor Planning Zones

Conceptually, the above is a good generic floor plan for many PCBs. To the far left side are sensors and devices with analog, small amplitude or noise

sensitive issues. Protective shielding and grounding strategies are used to isolate these devices. To the far right, are off board digital data streams with noisy surge currents due to driving external loads and cables. Also on the right is a HF MCU with switching noise. The power system is placed in the middle. Power can be noisy, but less problematic than digital switching noise. Additional power filtering may be needed entering the AFE region, and signal buffering between AFE and MCU helps keep noise out of the AFE.

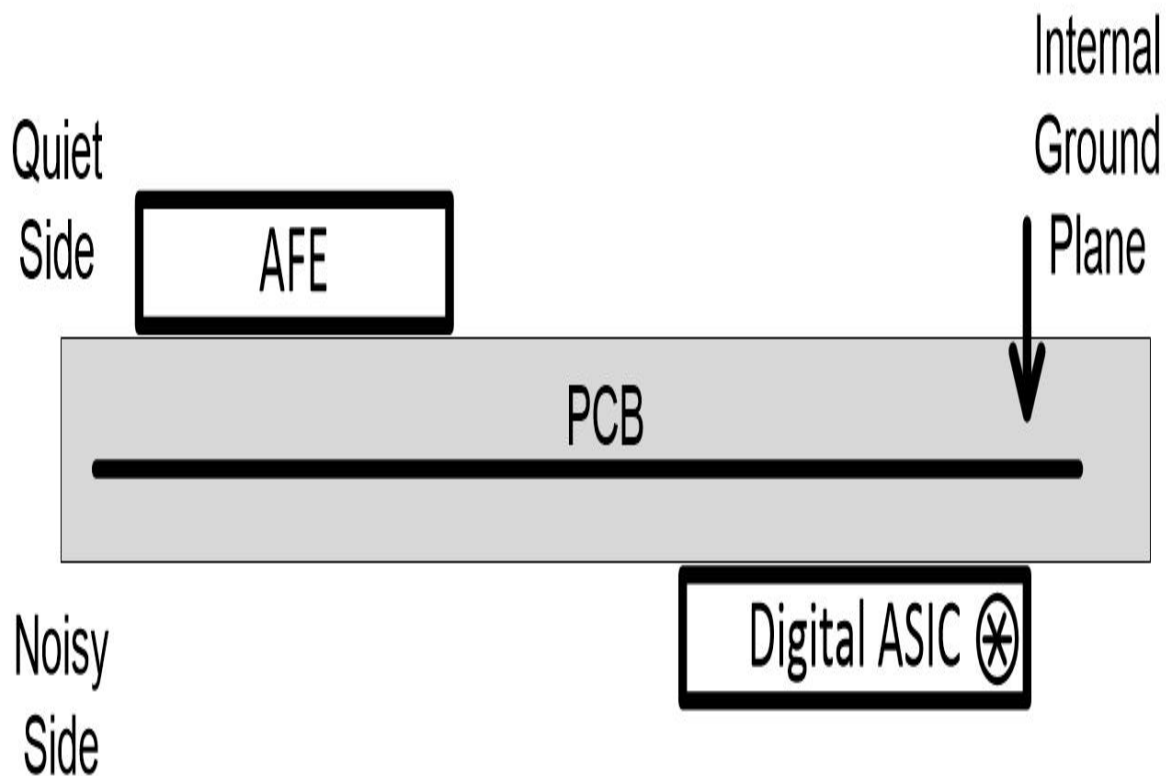


Figure 6-29. Two Sided Strategy

When putting components on both sides of the PCB a floor plan with a noisy side (digital), and a quiet side (noise sensitive) is often used. An internal ground plane serves as a shield between the two sides. This is a common strategy used in cell phones.

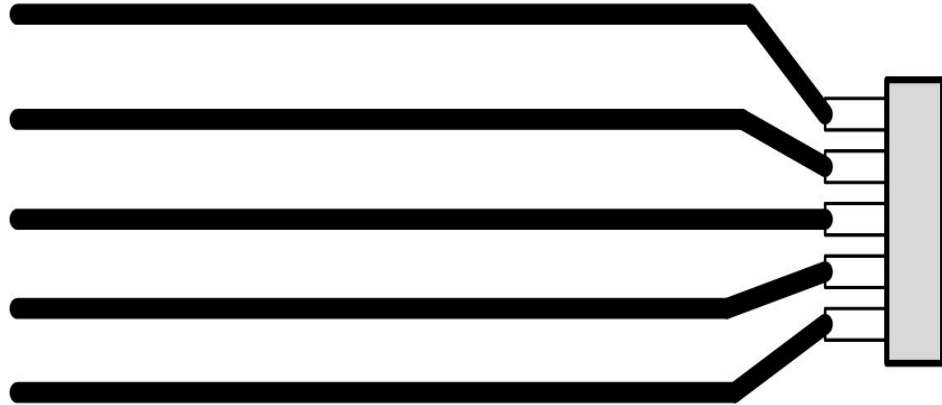
Floor planning to keep sources and victims apart helps greatly with noise performance. Selective placement of grounds and shields also reduce noise

coupling. Signal routing strategies are easily understood as seen in Figure 6-?.

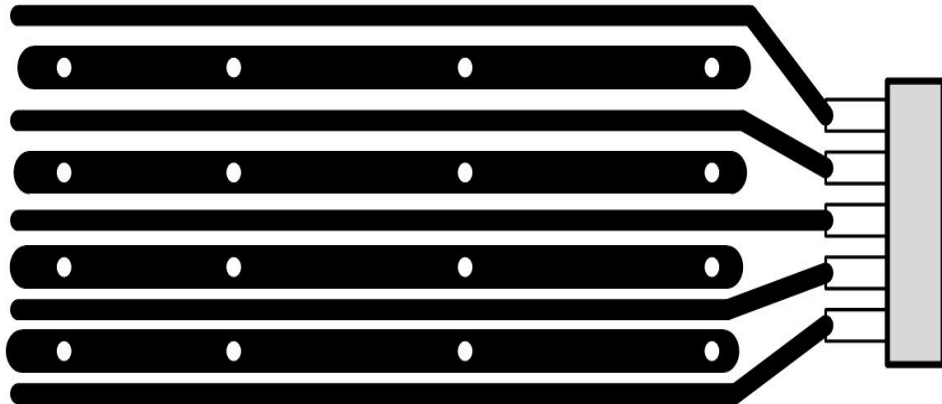
Case 1



Case 2



Case 3



Case 4

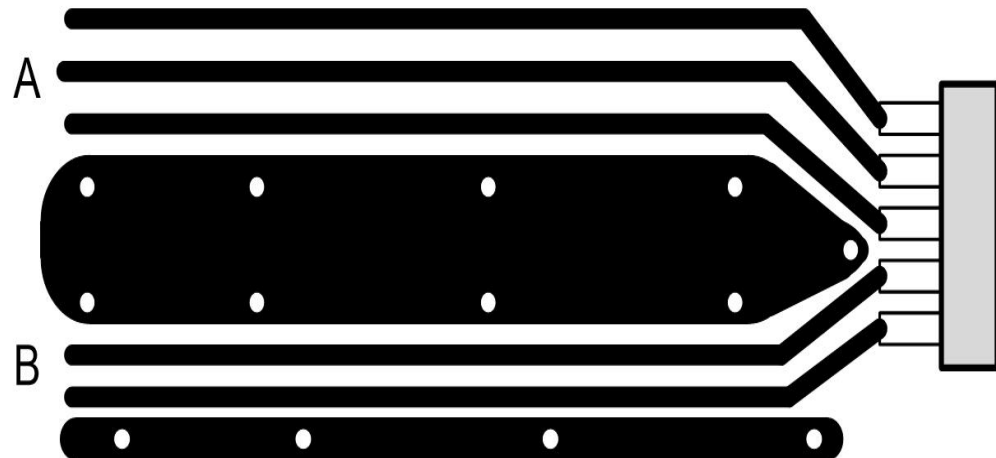


Figure 6-30. Reduce Coupling, Separation and Signal Splitting with Grounds

Case 1 shows 5 signals coming into the pins on a chip. The connections are made but there's no thought to the signal placement, and long tightly spaced connections often have capacitance and magnetic coupling issues.

Case 2 spreads the signals apart which should reduce the parasitic coupling between signals.

Case 3 introduces grounded conductors between the signals so most parasitic capacitance between connections are to ground and not another signal. Selective use of grounds between signals is a useful isolation technique. Vias are used to connect the grounding strips to the internal ground plane of the PCB.

The above three cases are a blind application of separation and grounding ideas. The optimal results include understanding the functional use of the connections first.

Case 4 determines connection functionality before routing. For this example, the top three signals (group A) are low activity digital lines and the bottom two (group B) are analog inputs of a differential ADC. The group A signals are not critical as either source or victim. Group B signals need to be routed as a differential pair for common mode purposes and be protected from possible noise sources.

A routing strategy that looks into the specific needs of signals can optimize placement and noise abatement methods as needed. (FDI:PCB)

A Faraday Cage (FC) is most often used for noise containment of entire systems. However, FCs can be used for localized containment of noise sources and selective protection of sensitive circuits within a system (Figure 6-?).

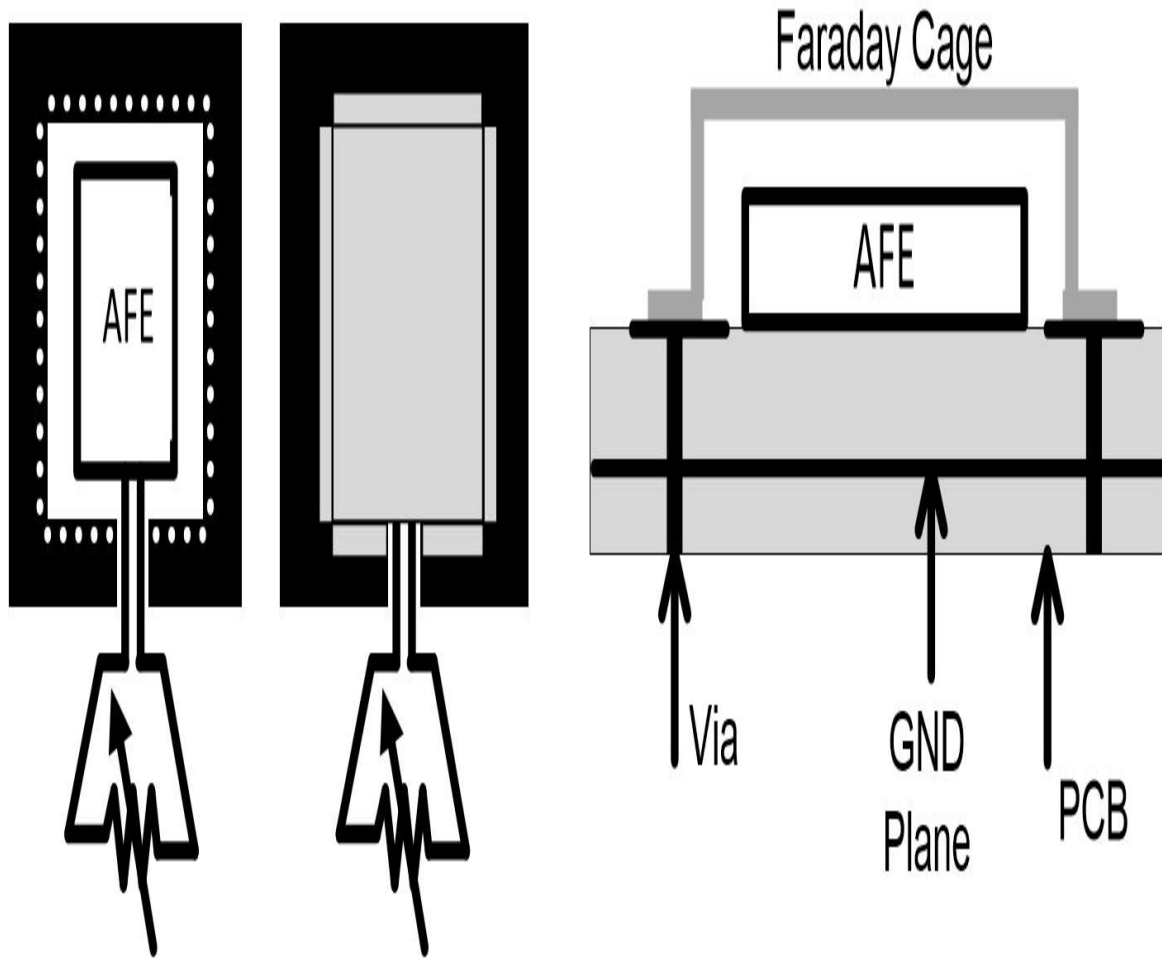


Figure 6-31. Faraday Cages to Protect AFE and Silence Noisemaker

This example shows use of a “metal shield can” that solders onto the PCB protecting an AFE. To perform properly, the PCB includes a solid metal ground plane underneath to fully enclose the device.

Multiple shield structures like this are common in most modern cell phones. Shielding around both sources and victims are required in the tight proximity of a cell phone.

Make Circuits Less Noise Sensitive

Inevitably, some noise couples to circuits. In all cases it is wise to take steps that make any devices less responsive to undesired signals.

Connection impedance can directly lead to undesired coupling (Figure 6-32).

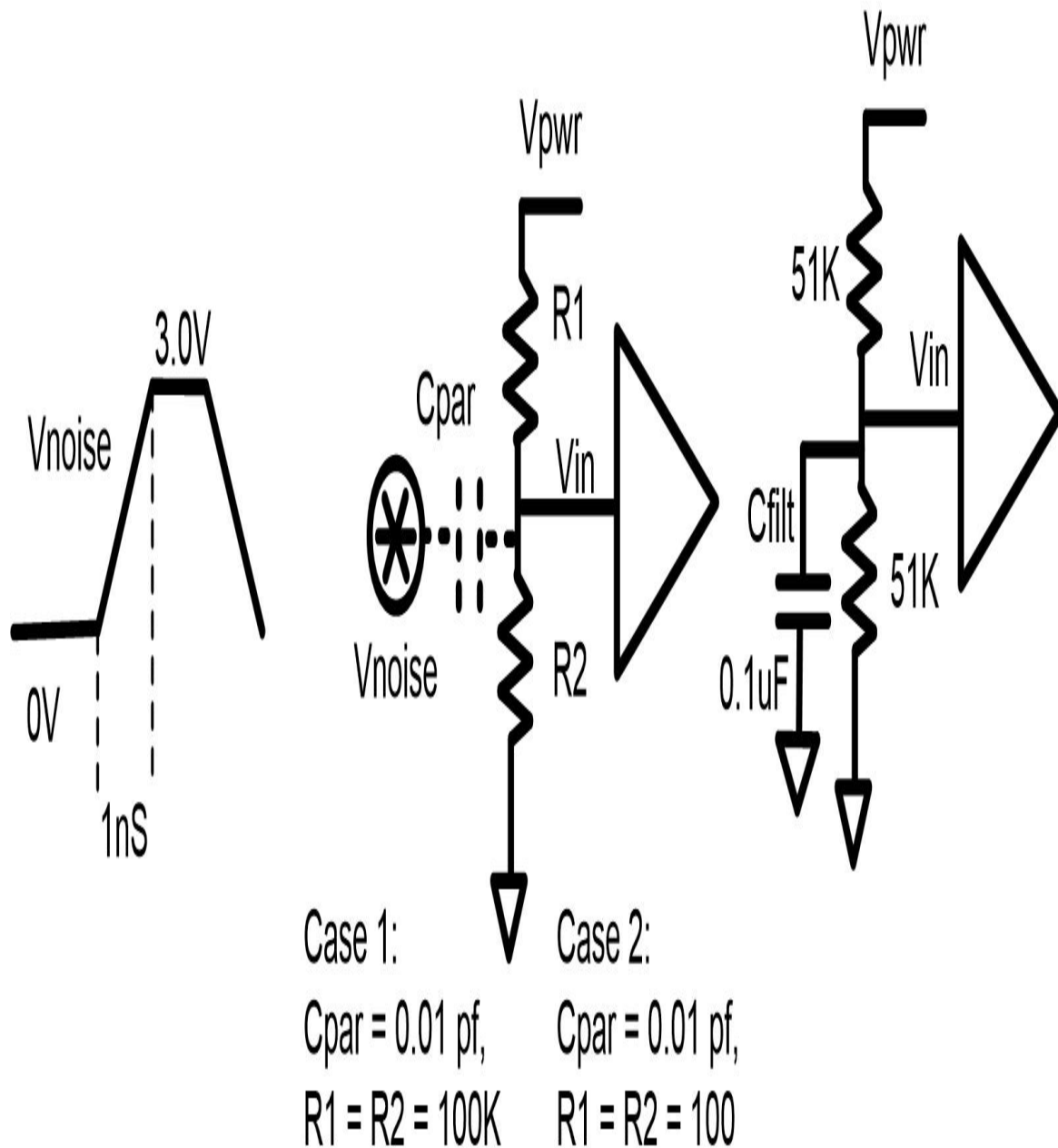


Figure 6-32. Weakness of High Impedance Nodes

In this example (above left), a voltage divider defines V_{in} . For this example, $C_{par} = 0.01$ pf is assumed, and is fixed as a function of the physical layout. As shown, the noise source is a digital signal with 3Vpp amplitude and 1nS rise-fall times. The transient edge injects ~ 30 uA of

current into the resistors. If $R1 = R2 = 100K$ ohms that 30 μA would result in a $\sim 3V$ noise transient. If $R1 = R2 = 100$ ohms the transient would be ~ 3 mV. With large resistors, noise can become a significant part of V_{in} .

Adding a filter capacitor (C_{filt}) makes a drastic change. Using a 0.1 μf filter capacitor creates voltage division between the two capacitors and the noise transient is reduced to microvolts.

A practical solution is shown on the right, with $R1$, $R2$, C_{filt} suitably sized. Adding C_{filt} has one additional improvement: noise on V_{pwr} is filtered for a more stable and quieter reference at V_{in} . It is important to locally filter any DC voltage. Small parasitic capacitors are common and this illustrates the importance of the issue.

Most RF front ends are done with an impedance near 50 ohms, keeping thermal noise low and suitable for transmission lines. That low impedance does consume current. For low frequency (<1 MHz) designs, staying within 100K to 10K is usually a good compromise on noise performance and power consumption.

Differential signals have advantages for both analog and digital. LVDS for high speed digital has been discussed, but more information for analog signals is useful (Figure 6-?).

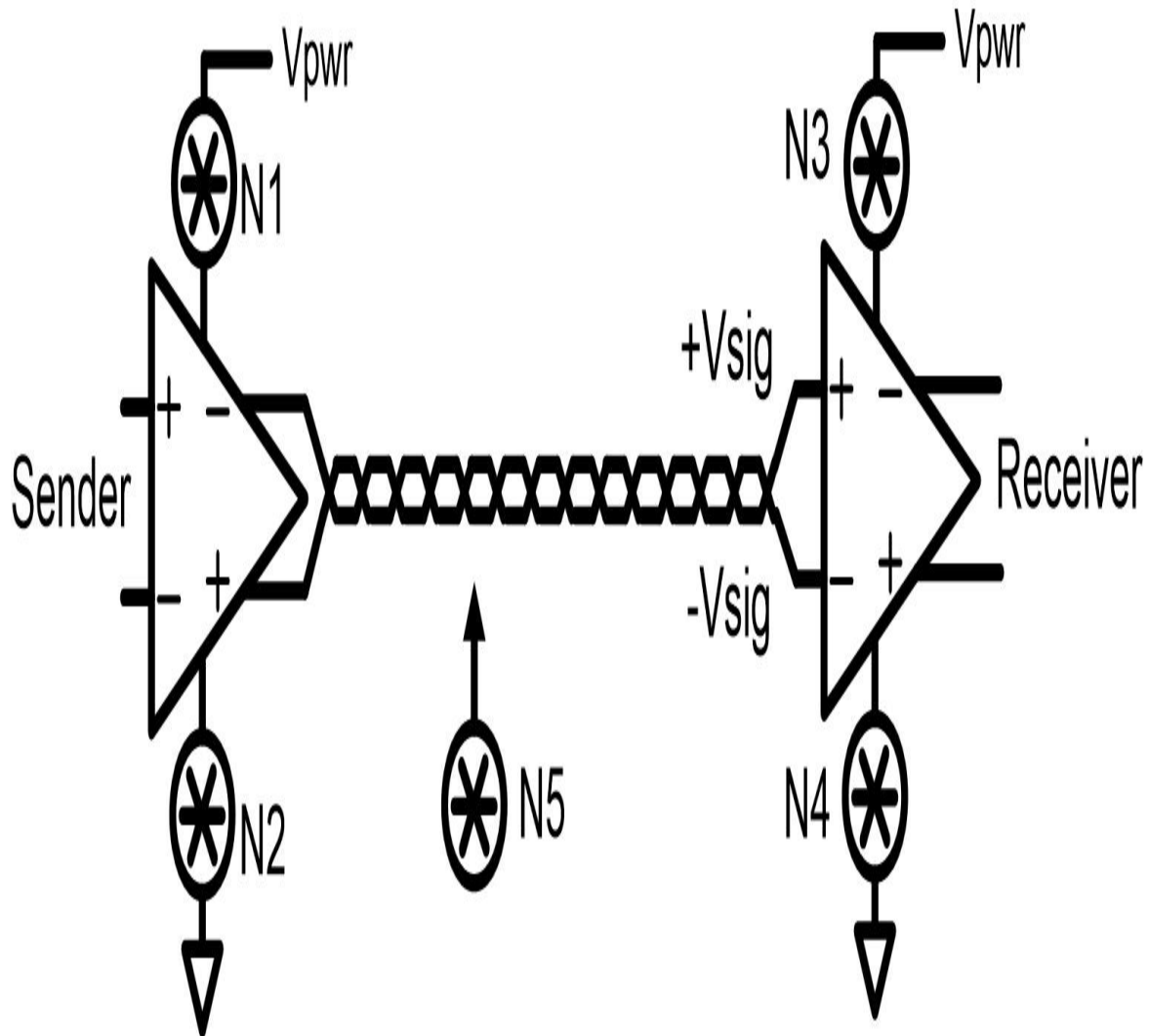


Figure 6-33. Differential Signals Reject Noise

As with digital LVDS, differential analog signals are defined by the difference between two signals (V_{sig+} , V_{sig-}) and are ideally independent from ground. The analog differential device is commonly called either a “Differential Output Instrumentation Amplifier” or a “Fully Differential Amplifier.” In audio designs, these are also referred to as “Balanced Amplifiers.” All of these can be thought of as op-amps that are differential at both input and output. Differential amplifiers reject common mode signals.

Consider the diagram shown. P&G at both ends of this system will have noise. Noise at the sender ($N1$, $N2$) may create common mode noise on the

sender output but are rejected by the receiver. Any noise coupled into the closely spaced interconnect (N5) is common mode and also rejected. Noise present at the receiver P&G (N3, N4) is also rejected by the receiver circuits as common mode noise. With better noise performance, the amplitude of differential signals can be smaller while still meeting desired signal to noise requirements.

Response time and bandwidth limiting are useful tools to create circuits that only respond to the signal of interest (Figure 6-?).

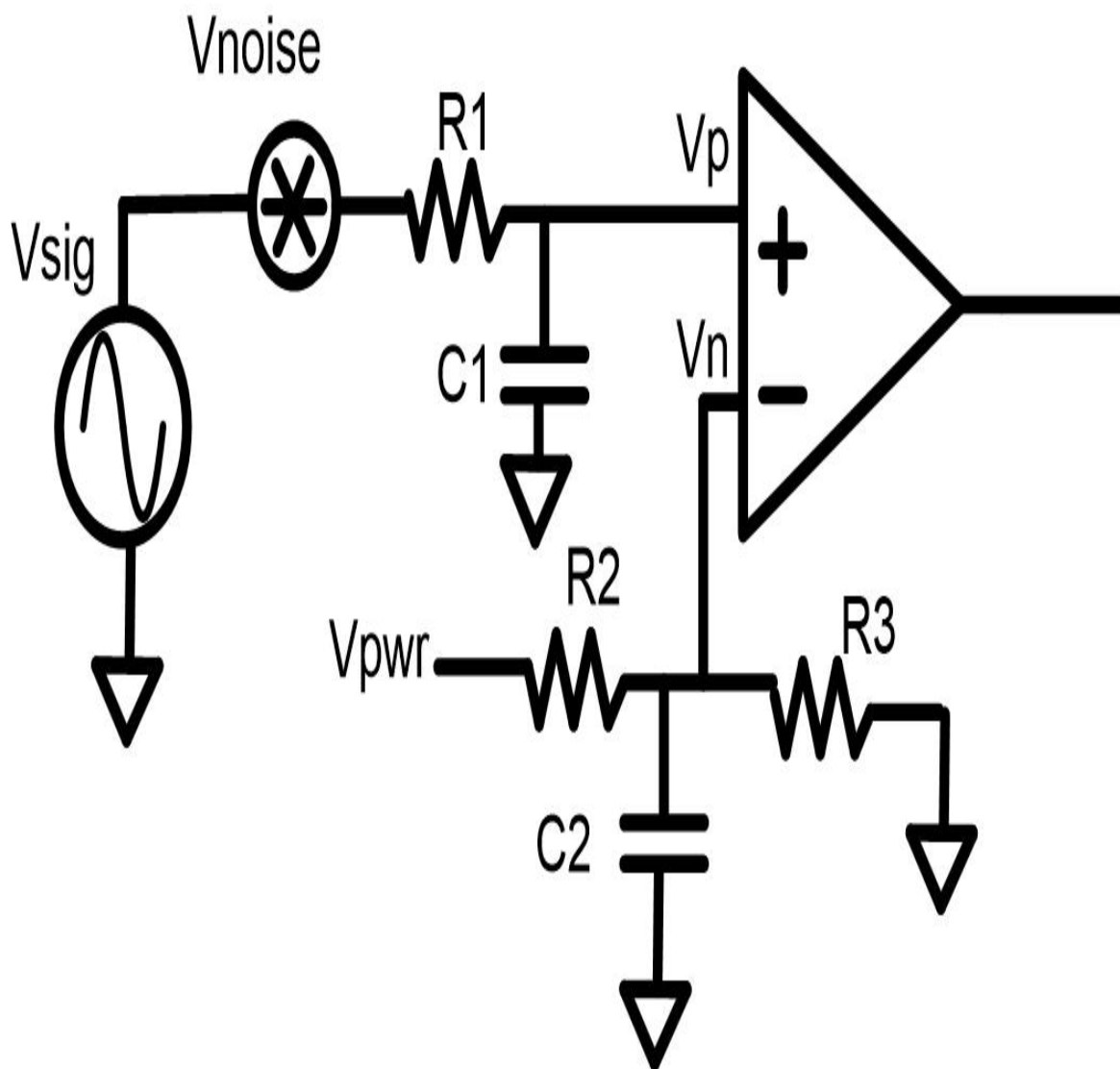


Figure 6-34. Bandwidth (BW) Limiting of Amplifiers and Signal Paths

If the amplifier in use has excess BW beyond the signal of interest, then limiting the input BW should be considered. Here, R1, C1 helps exclude noise beyond the BW of V_{sig} , and R2, R3, C2 establish a near DC BW for the reference voltage (V_n) applied to the comparator.

The response time shown below illustrates a fast vs. slow response to a noisy input signal. Response time and input filtering are commonly used tools to minimize false positives due to noise (Figure 6-?).

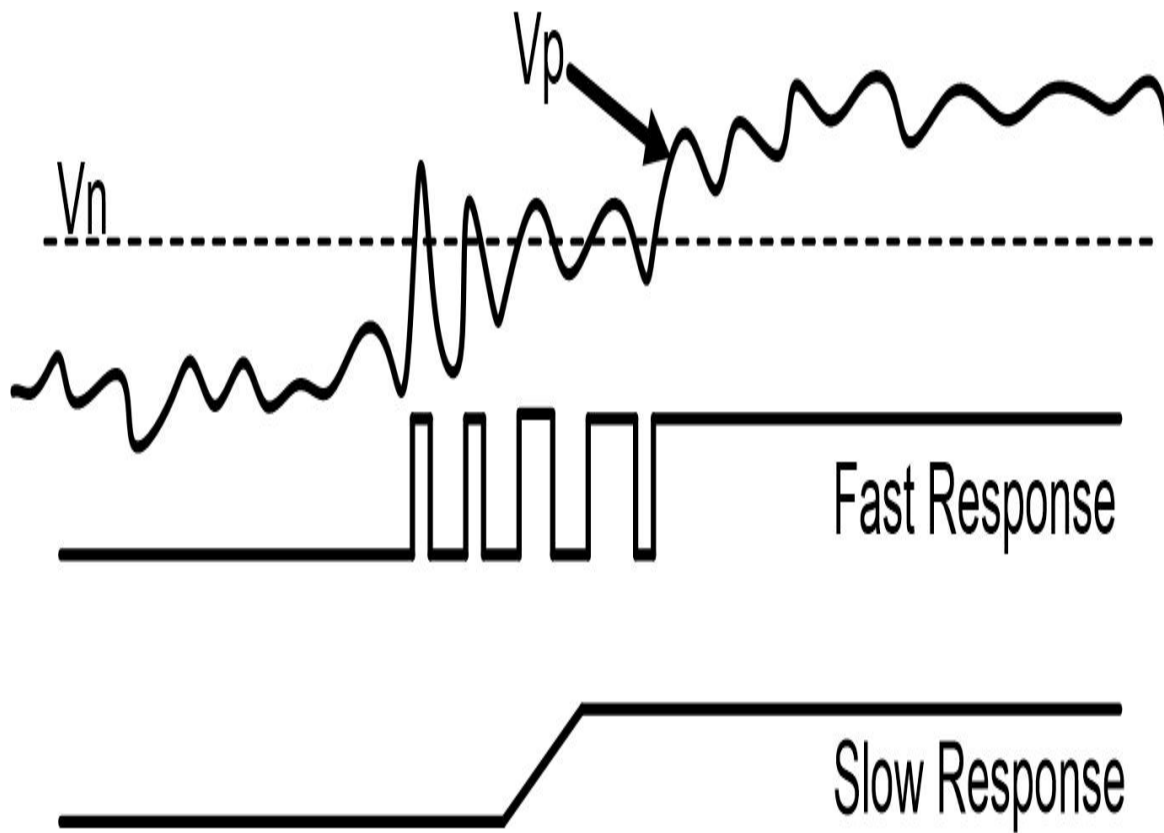


Figure 6-35. Comparator Response Time

This illustration shows a slow transient signal with imposed noise. The bandwidth (or response time) needed is defined by the signal. As an example, if V_{sig} is a tank level monitor and the tank needs 10 minutes to empty, there is no need for a circuit that responds in nanoseconds. Consequently, a slow response comparator and/or input BW limiting can avoid erroneous responses to high frequency noise.

Circuit response needs “just enough” speed for the signal of interest. Needlessly fast system electronics can generate false positives due to noise. Digital methods are also available through techniques of digital averaging or time window methods to determine if a stable state is present before recognizing a state change. (More: CODE)

Properly configured comparators create good noise immunity (Figure 6-?).

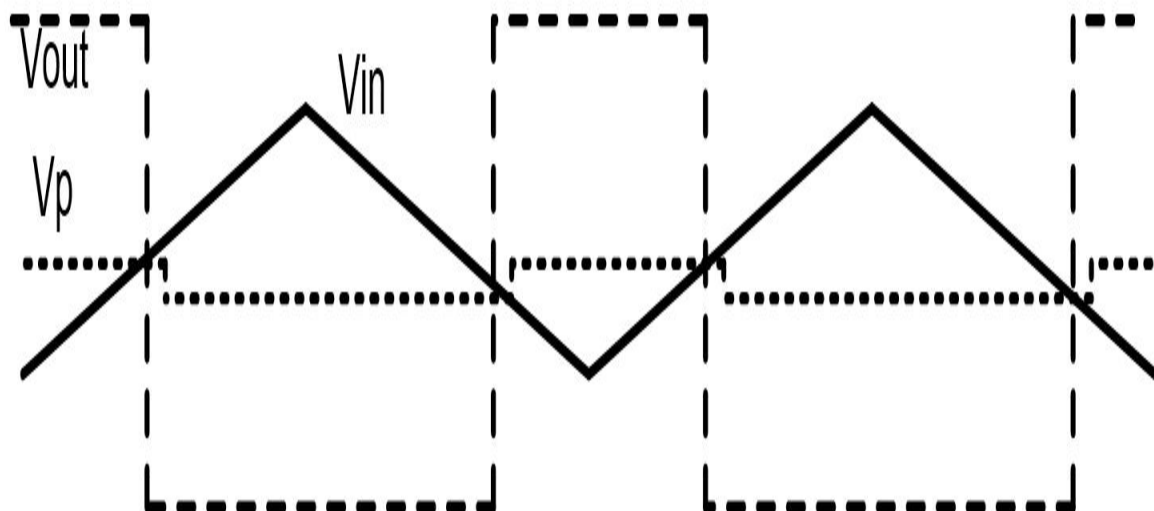
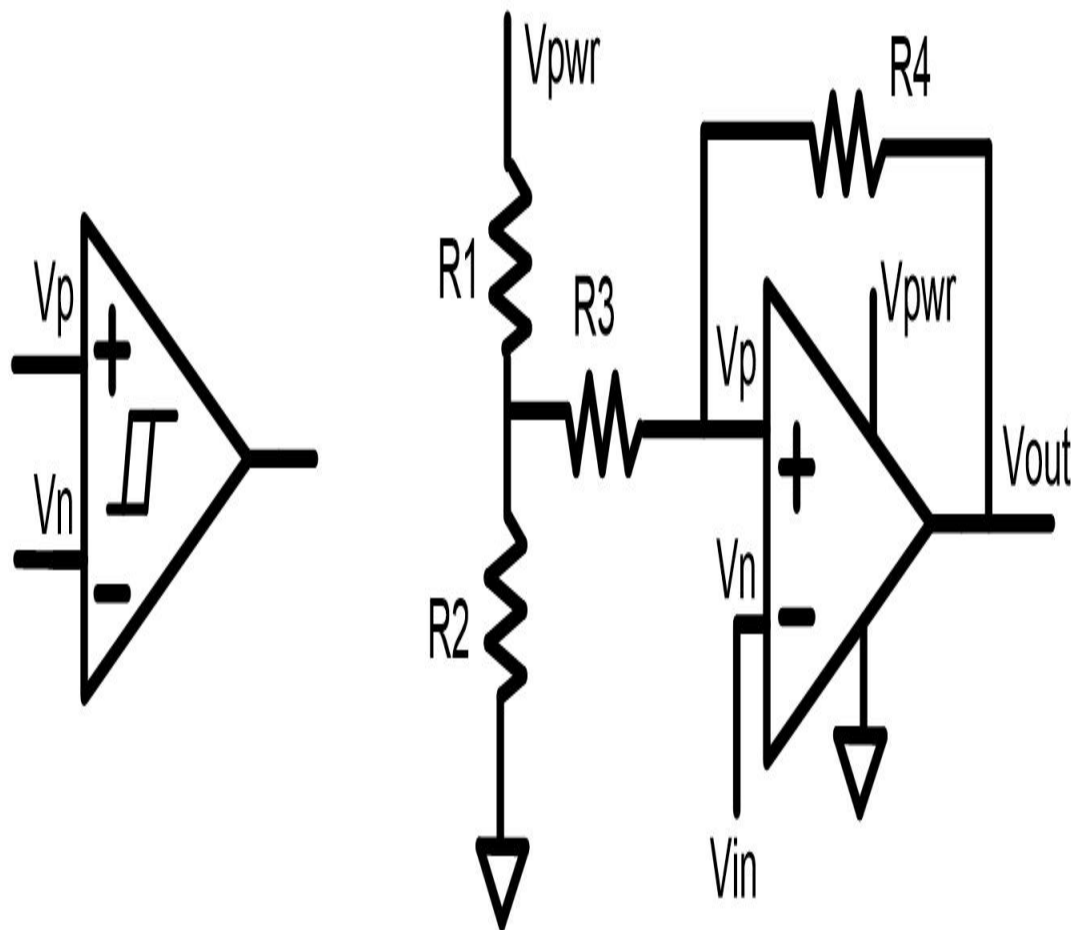


Figure 6-36. Hysterisis in Comparators

Hysteresis can shift the comparator switching point as it changes state. This is done internally using a Schmitt Trigger circuit, or done with feedback around the comparator that slightly shifts the reference voltage.

Non ideal comparator characteristics can also contribute to noise sensitivity (Figure 6-?).

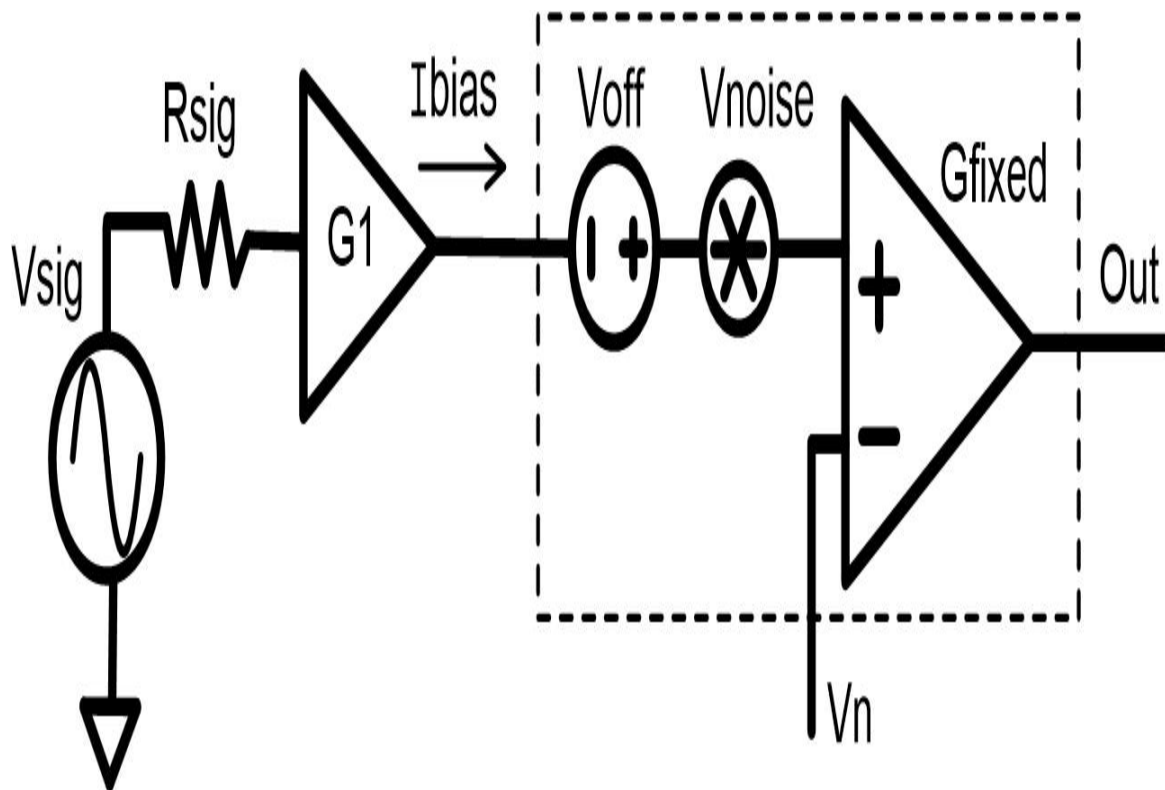


Figure 6-37. Non-Ideal Comparator

Consider:

- Small amplitude signals (V_{sig}) can be lost in the noise of the environment
- Input offset (V_{off}) affects the switching point. Offsets up to 10 mV are common, so the signal applied to the comparator needs to be much larger than the offset.
- Comparators have a fixed gain (60-100dB) and small amplitude signals may not create the full rail to rail digital output desired.

For signals, under 100 mVp-p, use of a gain amplifier might be needed between the signal and the comparator. The parameters of V_{sig} need to be compared to the performance specifications of the comparator to determine suitability.

Suppress Noise In-Out of System – Faraday Cage (FC) Techniques

Using a FC to partition off noise has already been introduced. Additional details are needed to implement a FC for an entire system (Figure 6-?).

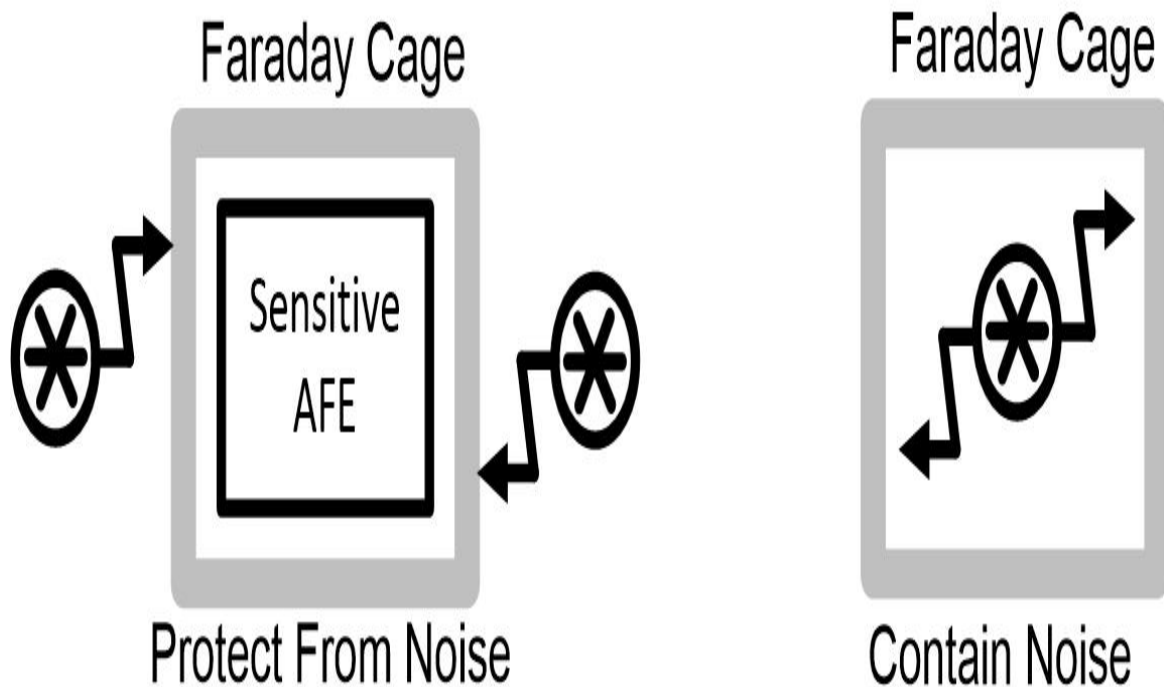


Figure 6-38. FC to Protect or Retain Noise

A FC can be used to protect circuitry from outside noise, or to contain a noise source. In some systems, multiple FCs are used to keep different parts of the internal system from interfering with each other. A FC for an entire system is frequently used to contain EMI and meet radiated emissions requirements.

When designing a system it can be unclear whether a FC is needed. Consequently, a good strategy is a design that easily adds FC capability

with minimal effort. The physical aspects of the FC need to be considered (Figure 6-?).

Plastic	Plastic W-CC	Aluminium	Steel
---------	-----------------	-----------	-------

Air Venting Without EMI Leakage

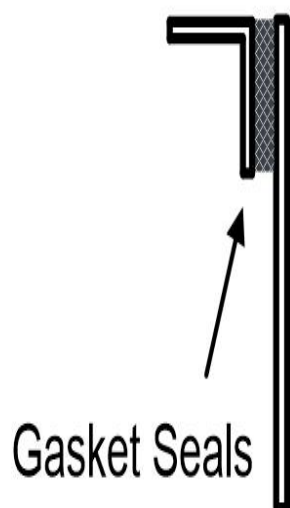
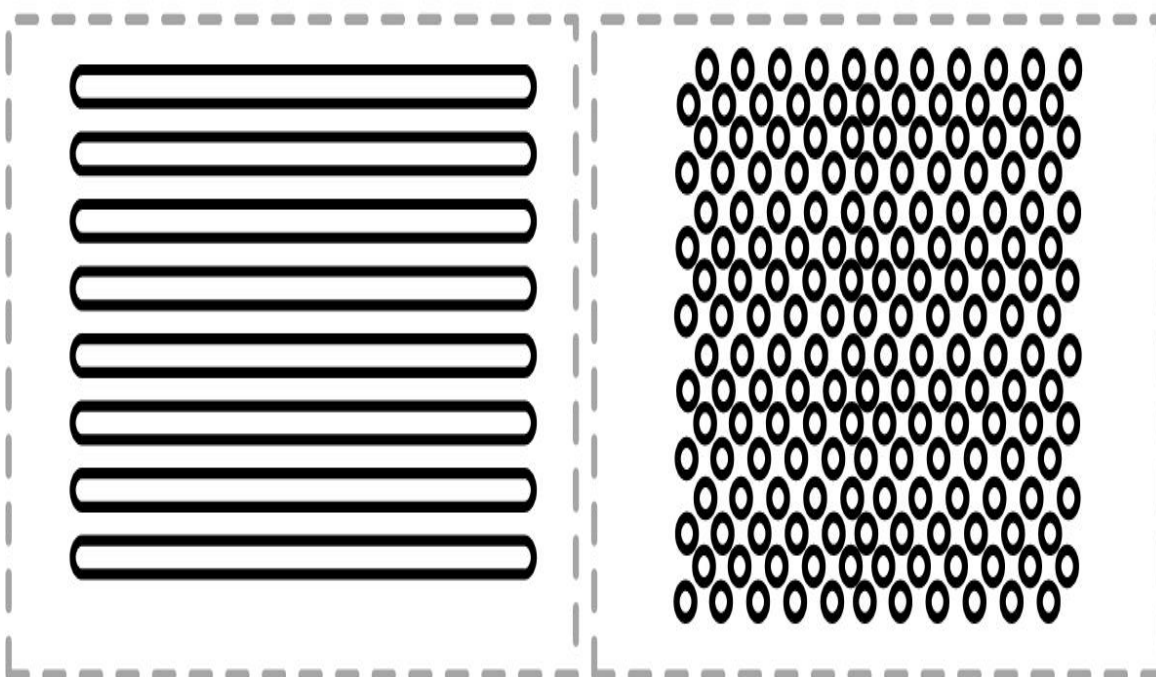


Figure 6-39. FC – Physical Considerations

The four common options for enclosure materials are:

- Plastic
- Plastic with internal Conductive Coating
- Aluminum
- Steel

Steel does best when containing or shielding magnetic fields. It is rarely a mandatory material however. Both aluminum and steel will contain radiated EMI when properly implemented. Plastic enclosures are low cost and easily implemented making them a popular solution. Plastic will not contain radiated EMI, but plastic enclosures that receive a conductive coating do well at reducing radiated EMI.

Devices put into a plastic enclosure should include the capability to easily add a conductive internal coating. In this manner, the low cost option of a plastic enclosure can easily shift strategy to a FC if radiated EMI requires it.

Air venting of an enclosure needs to be done without using slots, which can allow EMI to leak. Hole patterns are commonly used and work well. Air filters using fine metal mesh screens are also available for air passage without EMI leakage.

Enclosure seams need to be carefully thought out for proper sealing. Electrically conductive gasket materials are available to reduce EMI leakage at the seams. EMI absorbing materials are also available for lining the internal walls of an enclosure if necessary.

Connection ports going out of a FC can be problematic however. Some connection plugs were designed with EMI leakage in mind but most were not. Passing signals through the wall of a FC needs to keep the noise in while getting the signal out (Figure 6-?).

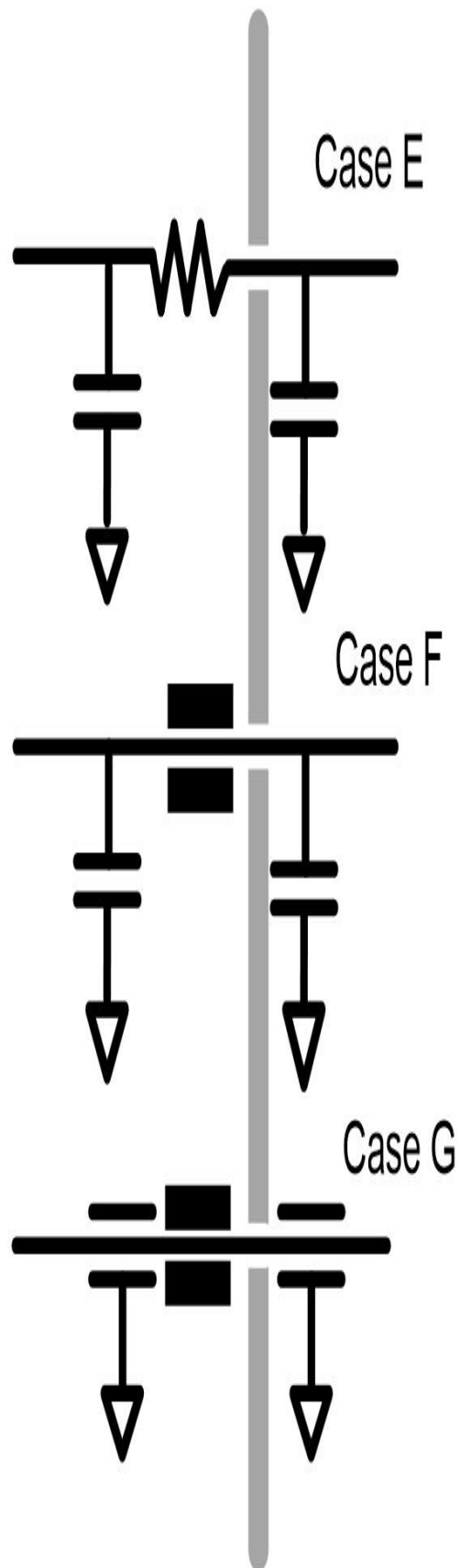
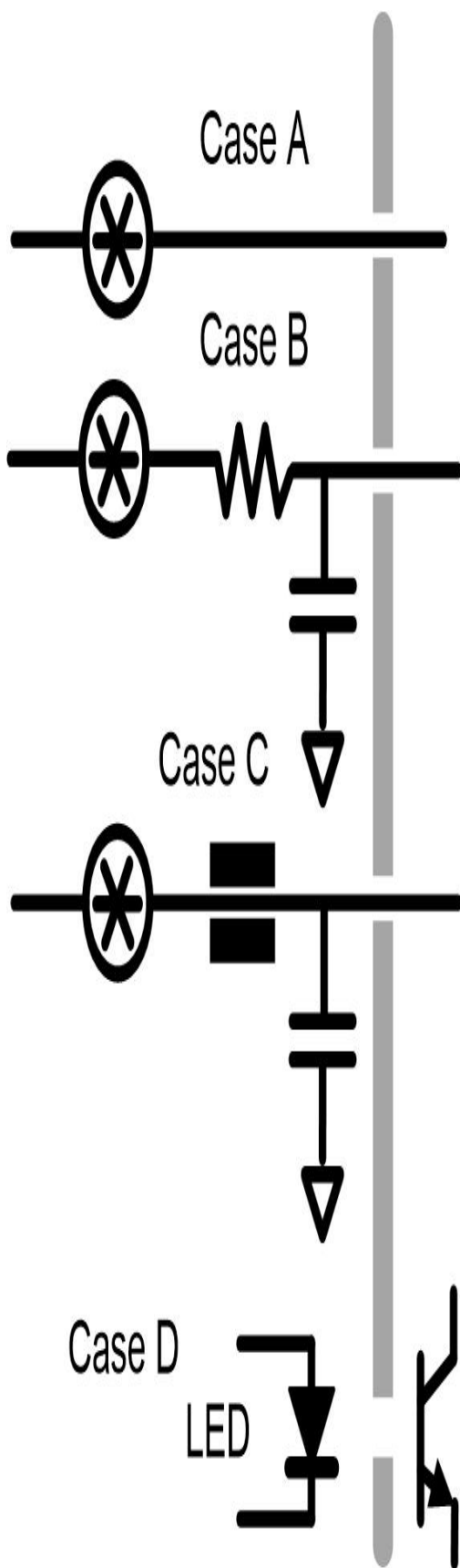


Figure 6-40. Connections In-Out of FC

Different methods of passing signals, involves different amounts of filtering:

- Case A shows a straight through connection, which can pass EMI on the wire to potential external antennas.
- Case B provides a LPF and is suitable for control signals that can tolerate some series resistance. Most digital control signals can utilize this with good results.
- Case C uses a FB instead, and is better suited to a current carrying connection.
- Case D uses optical isolation and is popular with high power switching power supply controllers.
- Case E&F shows filters that attenuate in both directions. Remember that noise can travel on connections in both directions.
- Case G uses “feed through” capacitors, which exhibit low series inductance (ESL) so that the filter remains effective at very high (typically > 1GHz) frequencies.

A suggested strategy for most situations uses Case E&F for PCB layout. A first PCB assembly with components can leave the capacitors off. Subsequent radiated EMI testing can determine if they are needed.

Electrostatic Discharge (ESD) Protection

ESD is commonly experienced when walking across a carpet and touching something conductive like a metal doorknob. Thunder and lightning are the most spectacular example of ESD. On a smaller scale, all systems need to survive ESD events undamaged, and be functional afterwards. ESD can destroy electrical circuits and designers need to deal with the over-voltage/current incidents that ESD can cause to their designs.

A system can be designed to both survive and keep functioning properly through an ESD event. Medical systems are required to do exactly that. Understanding how to do this starts with understanding a model of an ESD event (Figure 6-?).

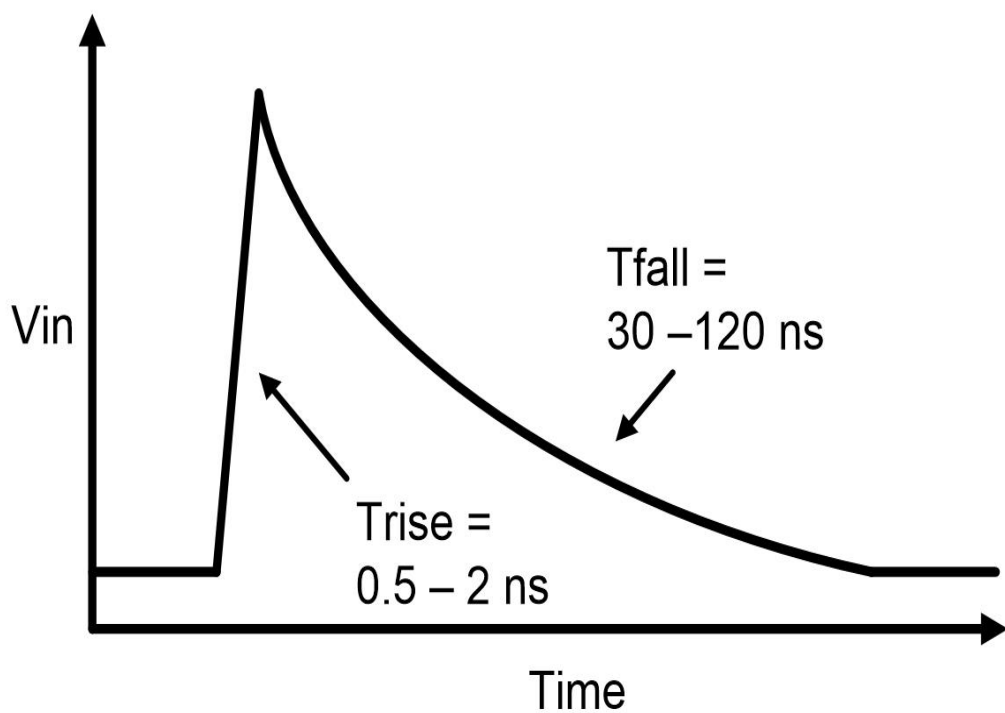
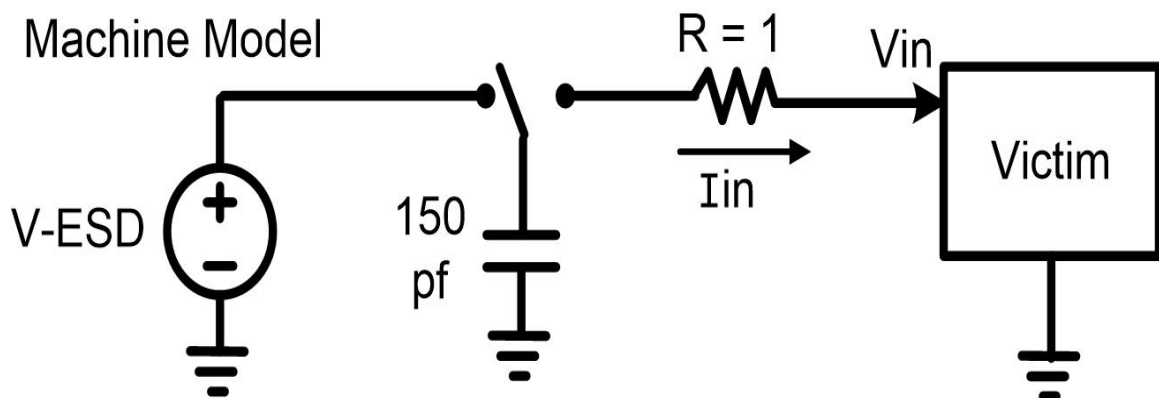
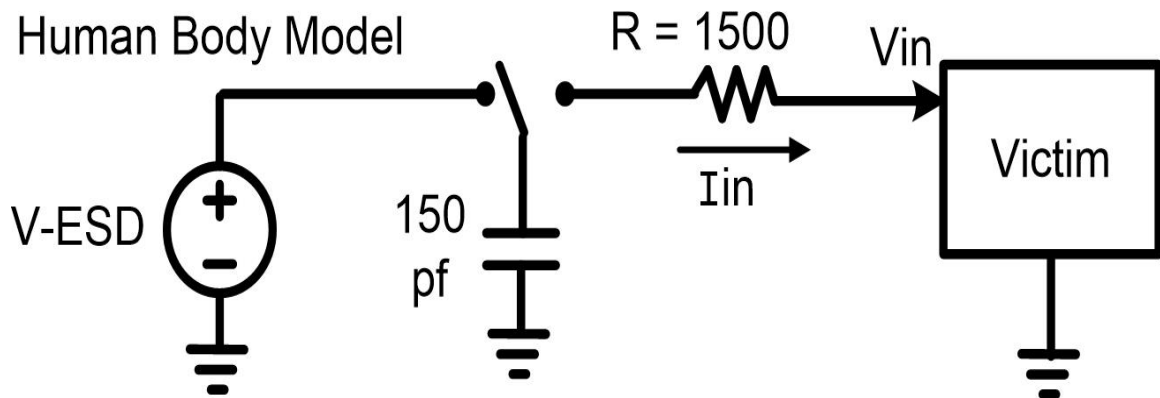


Figure 6-41. ESD as an Electrical Model Machine/Human Body Models

A simple model approximates a typical ESD event quite well:

- Charge a 150 pf capacitor up to V-ESD (-8KV to +8KV are commonly used)
- Connect the charged capacitor to the victim device
- A high voltage transient occurs at the victim

Typically, the voltage will transition to V-ESD in <2 nS and decay in <120 nS. ESD events can be applied anywhere to the exterior of a system. Anything conductive on the exterior of the enclosure is especially vulnerable. The charge polarity can be positive or negative so ESD protection needs to deal with both. Depending on the victim circuits, the sensitivity to a positive or negative strike can be very different.

The difference between the two most common models, the human body model (HBM) and the machine model (MM), is the series resistance. Human bodies (1500 ohms) don't conduct as well as metal machines (100 ohms).

The best protection against overvoltage damage is voltage limiting or clamping with a nonlinear circuit. Specialty transient voltage suppressor (TVS) diodes are commonly used. TVS devices have low impedance when forward biased or in their avalanche breakdown region (Figure 6-?).

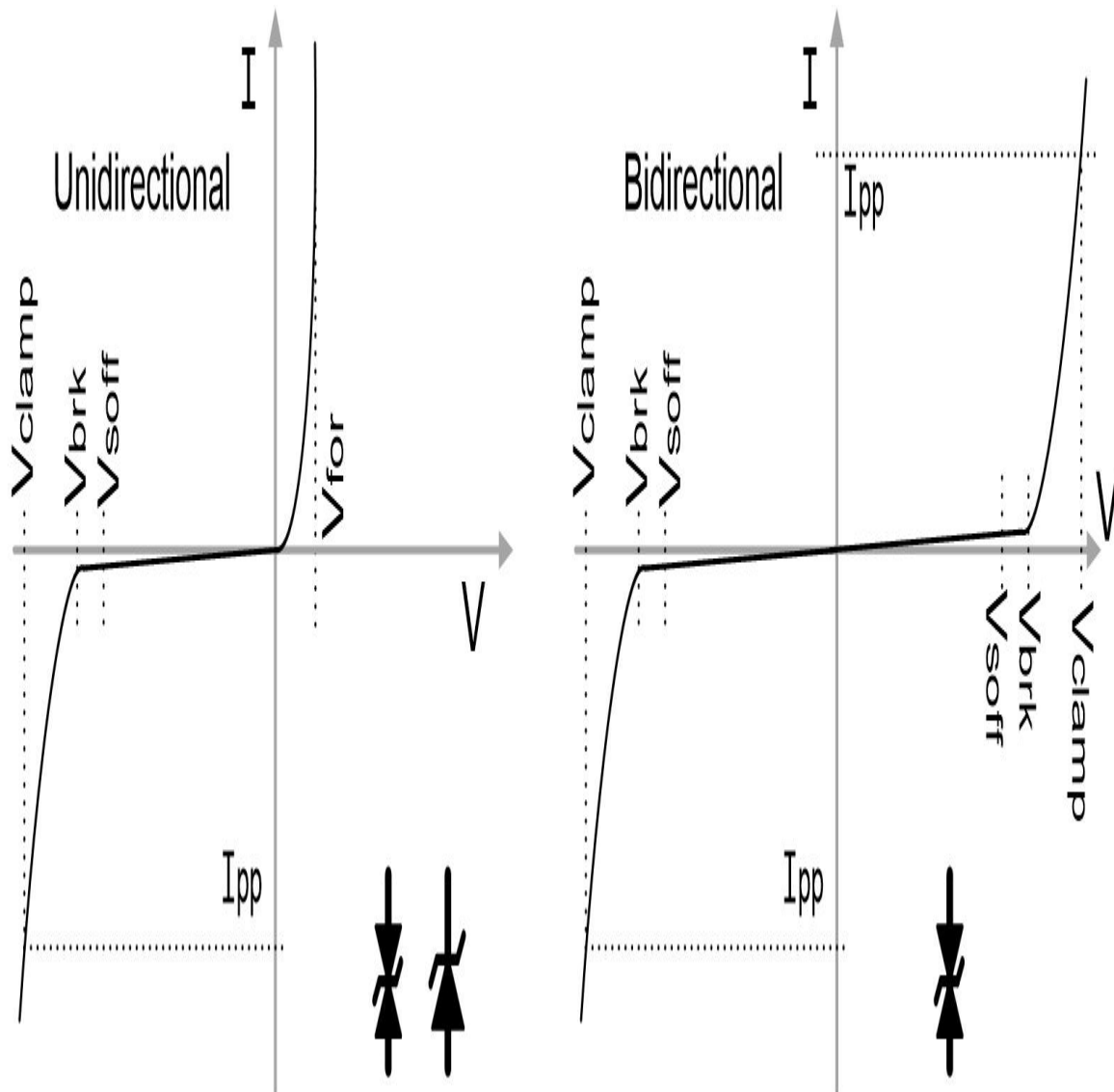


Figure 6-42. TVS Device Characteristics

TVS diodes function as voltage limiters. The TVS is optimized for high current transients over brief time periods. There are two common variants. Unidirectional TVS devices perform similar to a high current Zener diode. When forward biased, voltage and current are similar to a conventional silicon diode. When reverse biased (V_{for}), the voltage will be limited to a specified clamping voltage (V_{clamp}). In addition to a clamping voltage specification, a stand-off voltage (V_{soff}) is also specified. The stand-off voltage is the maximum voltage where the device does not turn on. As an

example, a 3.3V logic input could be protected using a Unidirectional TVS with 3.6V standoff voltage and 4.5V clamping voltage.

A Bidirectional TVS limits voltage equally in both directions. The “stand-off voltage” (V_{soff}) is the same for both positive and negative inputs. Most ESD protection is done using Unidirectional devices.

The stand-off voltages are available in values optimized to protect common logic levels and power supplies. One size does not fit all, and the TVS should have the appropriate stand-off and clamping voltages for the device being protected.

Any port on a digital IC has a very limited functional voltage range. Outside the functional voltage range, permanent damage to the IC can occur. On-chip ESD is limited in capability. Digital devices with off board connections need protection to avoid getting burned out.

Introducing a TVS quickly causes something else, namely a large current surge through the TVS (Figure 6-?).

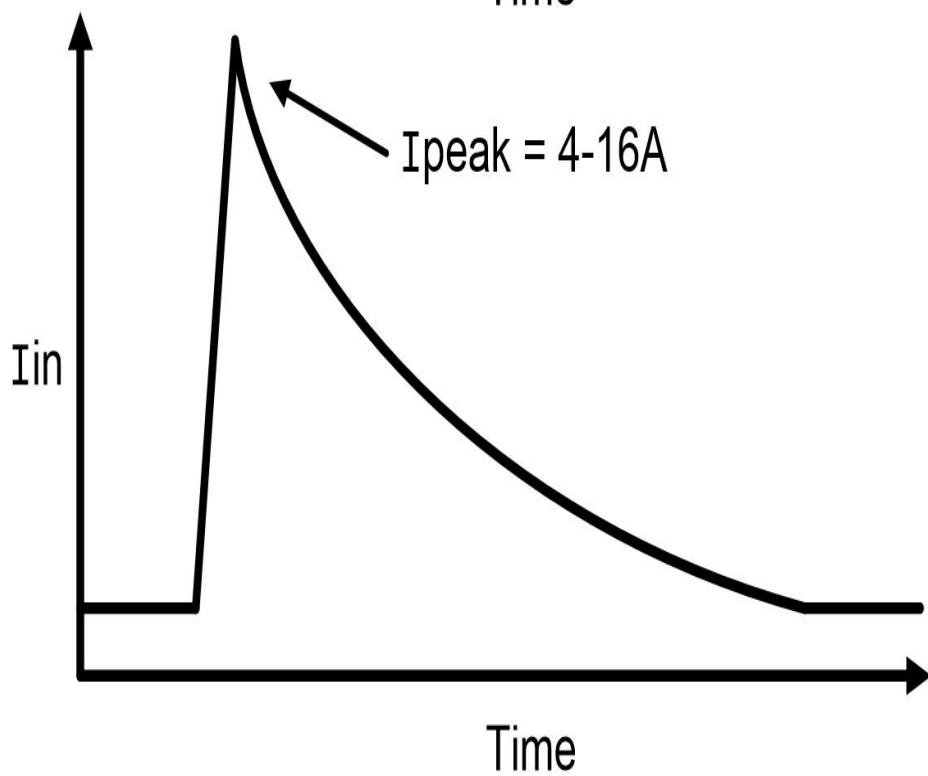
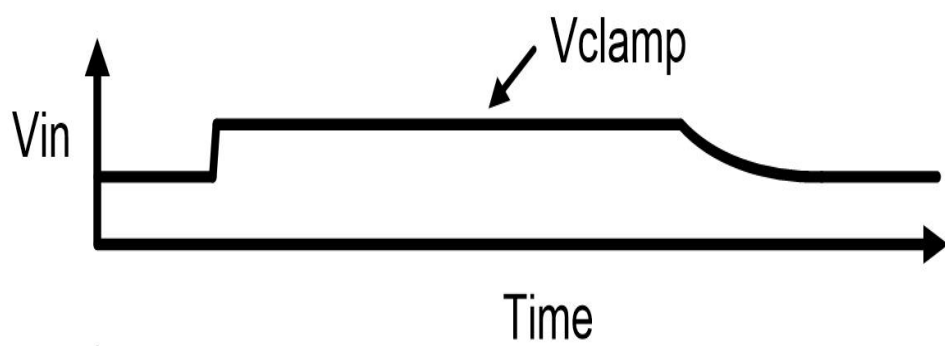
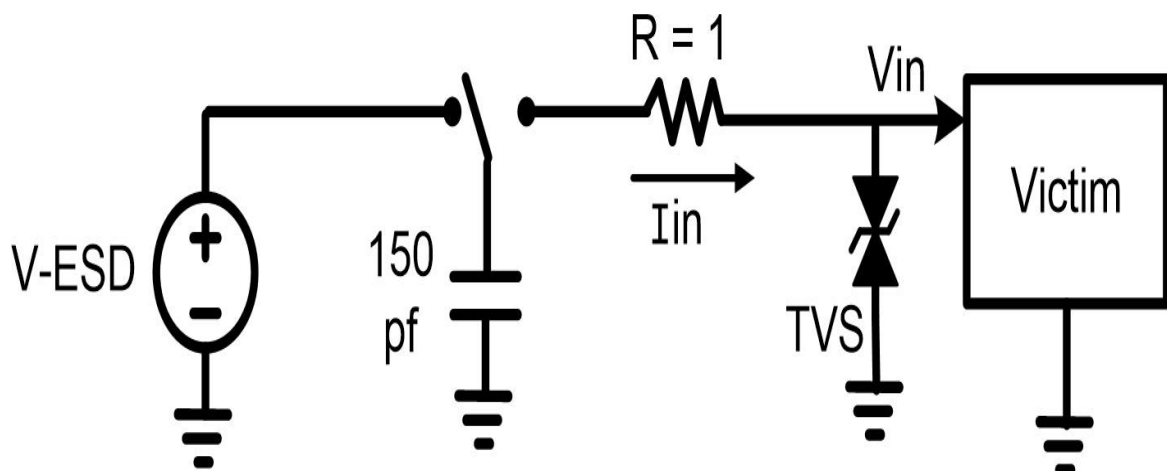


Figure 6-43. Injected Current Due to ESD Event

TVS use prevents overvoltage damage. But, the high voltage transient is replaced with a multi-amp current spike. That current spike can cause other problems. Depending on the total impedance of the connection, a 4-16 A surge current into the device is common. That huge transient current surge then becomes the newly created problem in ESD protection. Limiting the voltage is straightforward, but the resulting current can upset circuitry and grounds everywhere in the system.

Amps of current forced into the ground by the TVS causes inductive ringing of the system ground (Figure 6-?).

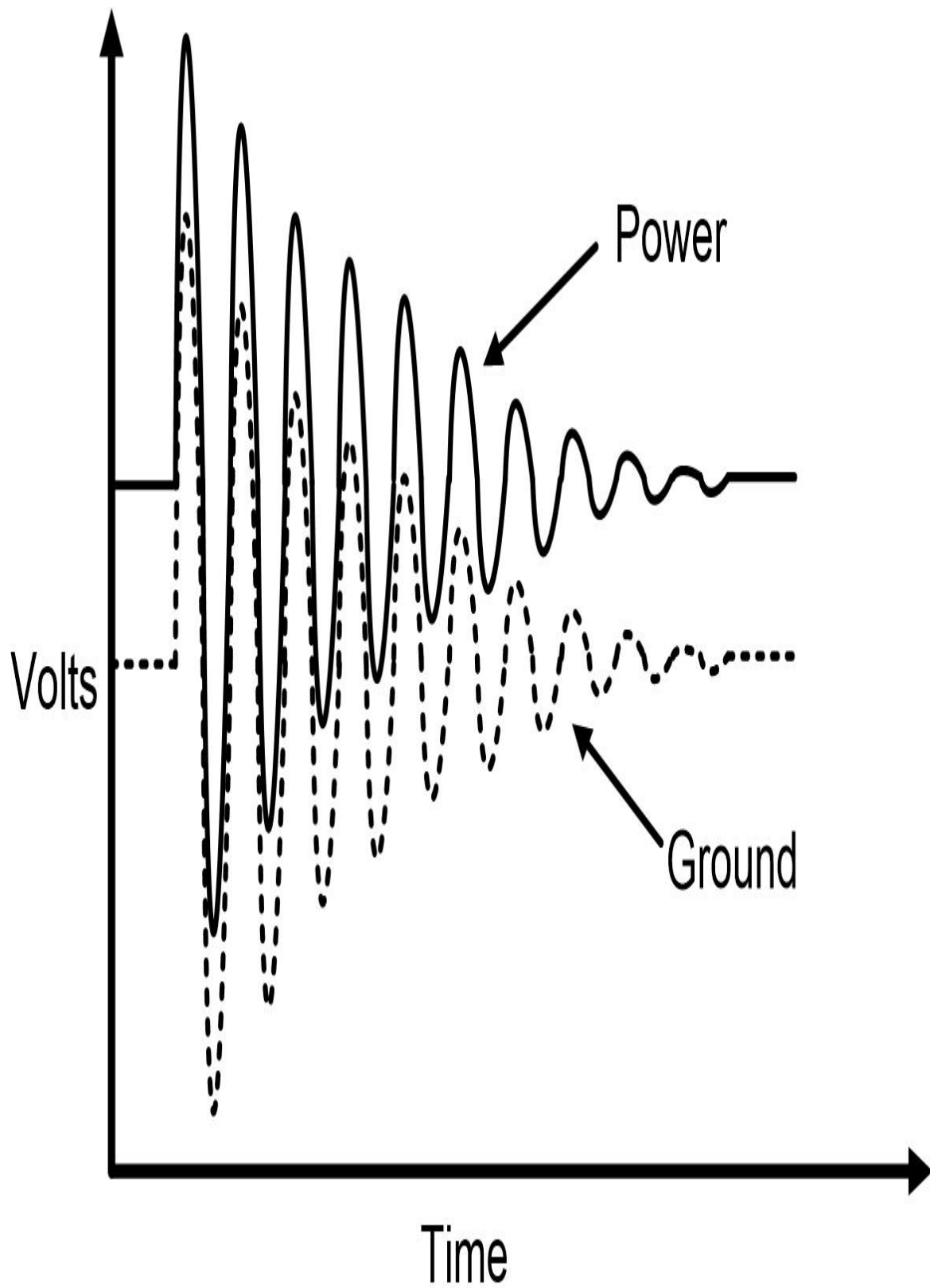


Figure 6-44. Ground and Power Bounce due to ESD

Injecting a large surge current into the ground through the voltage limiter will cause the PCB ground to bounce due to connection inductance. A board with a low impedance ground plane causes the entire ground to bounce together. Power connections, should have extensive high frequency decoupling to ground, and consequently the power bounces with the ground. With P&G bouncing together, proper functionality on the PCB should continue through the ESD event. Airplane turbulence and tightening your seat belt is a suitable analogy. Airplane or PCB, it will all stay functional if it all bounces together.

Establishing P&G integrity allows the PCB to maintain signal integrity across the board during an ESD event. However, external control lines are defined relative to an external ground, and can get corrupted. An ESD event in one location can cause an input elsewhere to be seen as faulty. The solutions for this are multiple:

- Inputs have to travel with the PCB ground, using a LPF to do so
- Inputs can be non responsive to fast transients, by software state sampling
- Inputs use an error detection/correction protocol

The elapsed time for the ESD event is under 150 nS for the initial current injection, and the ringing typically settles out in under 100 mS. Filtering, sampling, or error detection methods can readily deal with this.

Understanding what happens during an ESD event leads to the question of what needs to be protected from ESD (Figure 6-?).

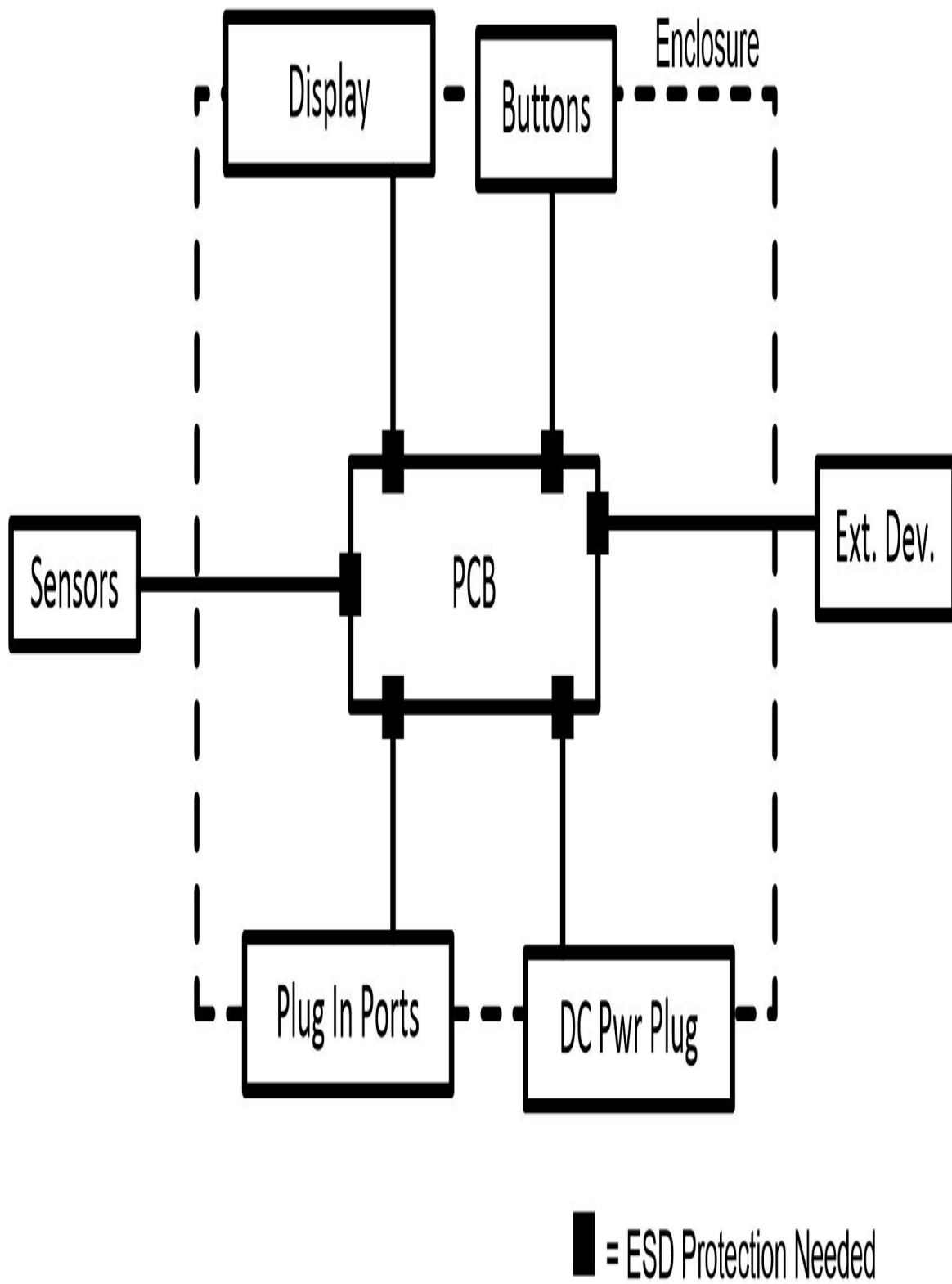


Figure 6-45. ESD Strategy as a Protective Fortress

Some form of protective enclosure will be part of the device. Internally isolating the PCB is the first line of defense that needs to be carefully thought out. The device enclosure becomes the figurative fortress. With board-level ESD, building a fortress and establishing protected access points “across the moat” becomes the strategy. All items connected out beyond the “castle walls” need to be ESD protected.

When defining the enclosure, remember that ESD to any external part of the case can take multiple paths into the electronics. If there is a conductive path of any form to the PCB it needs to be ESD protected as an external port, connected to ground, or removed from the conduction path.

Protection circuits should be placed at the entrance to the board and not downstream from the entrance point. ESD can be kilovolts and creates possible arcing problems. This is best dealt with immediately where the connection enters the PCB. Many possible options for protection exist (Figure 6-?).

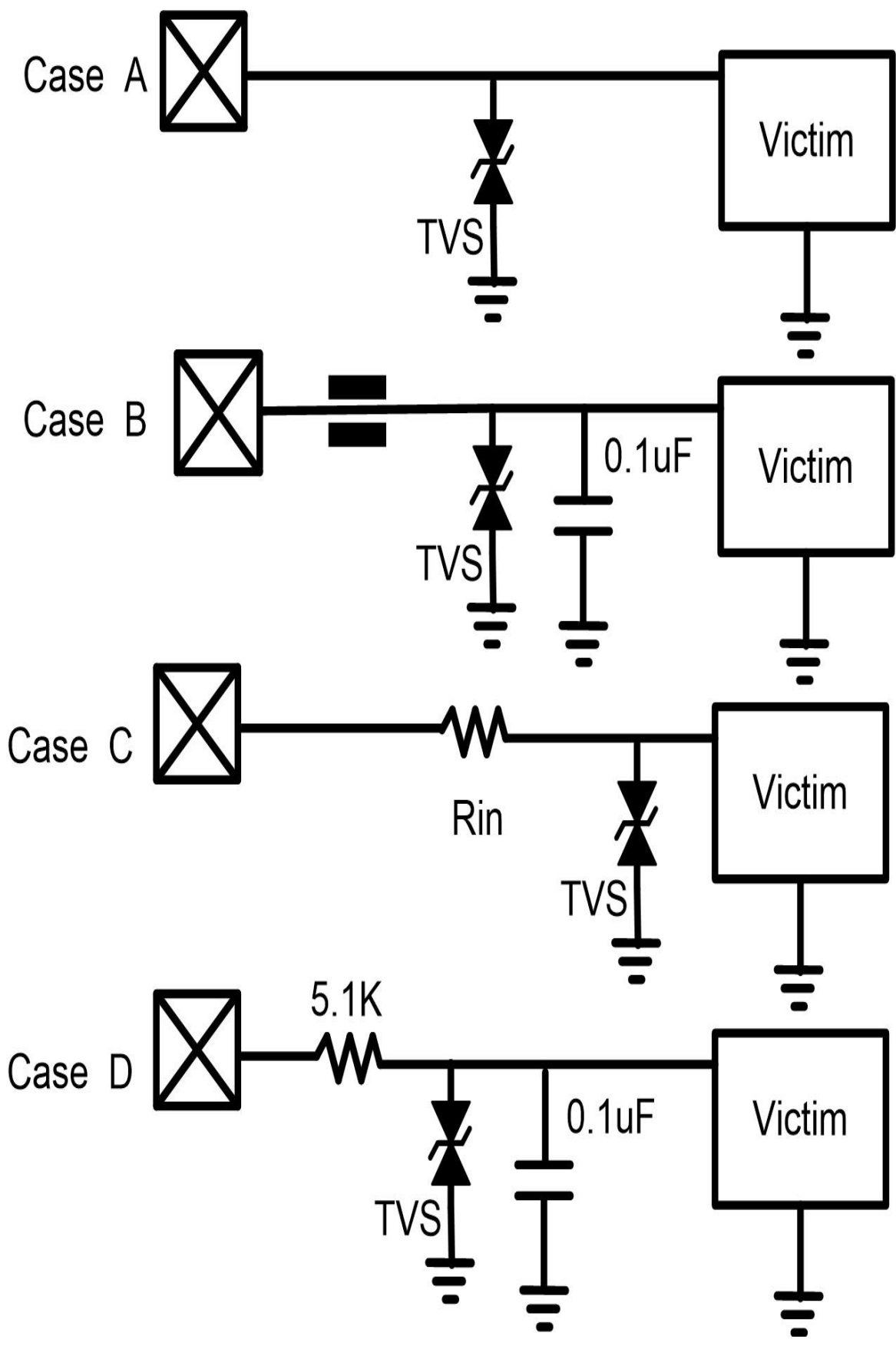


Figure 6-46. TVS Strategy with Supplemental R,L,C,FB Devices

Case A – A basic TVS limiting circuit provides overvoltage protection but creates surge currents into ground. For connections that can't tolerate any series impedance it is better than nothing, but other problems may emerge.

If possible, series impedance reduces the current injected into ground.

Case B – adds series impedance using a FB to the ESD signal path, and helps limit the surge current. The capacitor helps with high frequency stability and the connection is still low impedance at DC. This is the preferred method for power connections. The FB also prevents the power wire from creating a radiating antenna.

Case C adds series resistance to the connection which greatly reduces the current that gets pumped in the ground. For things like pushbuttons, low frequency digital data, significant (1K to 10K) series resistance can usually be tolerated. High speed data requires smaller resistance, but $R_{in} = 100$ on fast data connection is often tolerable and still helps reduce the ESD created current.

Case D adds filtering capacitance to keep the input to the victim to remain stable through the ground bounce of an ESD event. The RC time constant only allows slow (pushbutton type) responses, and can be made faster as the input might need more speed. This method works well to protect the large number of low-bandwidth inputs, which includes “set and forget” control lines, sensor inputs, and similar.

Although discussing input port protection here, output port protection is both similar and also needed. TVS and additive resistance are appropriate. Voltage limiting helps prevent damage to semiconductors and other parts with voltage restrictions.

For high speed data, a properly designed communication protocol with error checking needs to be used. ESD would corrupt data, and bad data prompts a request to resend. Since an ESD event, and aftereffects, lasts under 100 ms, the check, verify, and resend process deals with ESD invisibly. The end user is never aware that corrupt information was corrected. Some protocols don't

check and correct data. If data goes off the PCB it needs error check and correct capability. (More: DATA)

Essential Concepts – Conclusions and summary

EMI has to be considered as part of any electronic design. Noise can be both a regulatory and performance issue. Too much radiated noise can prevent device approval and commercial sale. Too much internal noise coupling can reduce performance & accuracy, or even render a device nonfunctional. Many things that can be done to mitigate noise problems don't add cost to the product, and should be included in any design.

A simplified model for noise coupling sees three things: source, medium, and victim. This model helps to understand how noise is conveyed between devices. Noise reduction techniques generally fit within these areas:

- Silence the Source
- Contain the Source
- Protect the Victim
- Deafen the Victim
- Separate the Source & Victim

An initial design should approach noise reduction as an application of best practices. Recognize both noise sources, and noise sensitive victims. Then, take actions to silence, separate, shield, or render non-responsive.

Frequently, that is sufficient. After implementation, noise testing is best observed in the frequency domain where the frequency of noise sources provides well defined clues to the circuit sources.

Transient switching is the predominant EMI source. Noise coupling paths include P&G connections, parasitic capacitance, inductive coupling, and radiated noise signals. Low impedance grounding serves as a foundation for

any EMI reduction strategy. But, no ground is ideal and small amounts of dynamic voltage variance in a ground are expected and need to be accounted for.

Open current loops should be avoided to minimize radiated EMI. Twisted wire pair interconnects, differential signal pairs, LVDS, and PCB connections over a ground plane all minimize current loops and their EMI.

Consider using slower clocks and slower transient edges for all digital signals and clocks. Frequently devices can run slower, saving power, reducing EMI, and avoiding the need for impedance matched transmission lines.

Any SMPS power regulator should include PCB locations for RC snubbers, ferrites to isolate the SMPS, and high frequency filter capacitors. Heat sinks should be grounded out, and motors should receive RC snubbers with ferrites on their cables. Any DC reference voltage should have decoupling capacitors in close proximity to where that DC voltage is used. Creation of independent power sources for analog and digital devices are greatly preferred.

Designers should understand, what parts are noise sensitive victims, and what parts are noise sources. Using that, implement a strategy that separates victims and sources. Liberal use of grounds between sources and victims aids in their separation.

Bandwidth limiting of amplifiers, comparators, and signal paths improve noise immunity. Lower impedance connections and bias circuits are less sensitive to capacitive crosstalk.

Faraday cages can be used to reduce noise from an entire system, protect sensitive internal devices from noise, or contain noise emitted from a particular circuit. Frequently multiple FC's are used within a single system to achieve needed noise performance. In all cases, signals in-out of the FC need filters so that noise does not leak via electrical connections.

ESD protection will be required on all external connection paths out of the device enclosure. For some products, ESD protection will be defined and

tested for as a regulatory requirement. For other devices, a lack of ESD protection becomes a liability as products fail in the field.

All ideas presented here on EMI and ESD are readily implemented into a design, many with minimal effort and expense. Reviewing schematic design, PCB layout, and enclosure implementation specifically for EMI & ESD issues should be done.

Further Reading

Controlling Radiated Emissions by Design, Michel Mardiguian, 1992, Chapman and Hall, ISBN 0-442-00946-6

Noise Reduction Techniques in Electronic Systems, Henry W Ott, 1976, Wiley-Interscience ISBN 0-471-65726-3

Electromagnetic Compatibility Engineering, Henry W. Ott, 2009, John Wiley & Sons, ISBN 978-0-470-18930-6

Signal and Power Integrity – Simplified 3rd Edition, Eric Bogatin, 2018, Pearson Education, Prentice Hall, ISBN-13 978-0-13-451341-6

EMC for Product Designers 2nd Edition, 1996, Tim Williams, Butterworth-Heinemann, ISBN 0-7506-2466-3

Common Mode Filter Design Guide Document 191-1, Revision 11/08/07
Retrieved from: Coilcraft.com

Simple Grounding Rules Yield Huge Rewards, Jerry Twomey, April 27, 2012, Electronic Design Magazine, Retrieved from:
<https://effectiveelectronics.com/articles-patents/>

Quiet Down to Meet FCC Emission Standards, Jerry Twomey, July 10, 2013, Electronic Design Magazine, Retrieved from:
<https://effectiveelectronics.com/articles-patents/>

Protect Your Fortress From ESD, Jerry Twomey, August 9, 2012 Electronic Design Magazine, Retrieved from: <https://effectiveelectronics.com/articles-patents/>

EMC and System ESD Design Guidelines for Board Layout, AP24026
Infineon Corp.

Retrieved from: <https://www.infineon.com/>

IEC 61000-4-2:2008 Testing and Measurement Techniques – Electrostatic
Discharge Immunity Test, International Electrotechnical Commission
<https://www.iec.ch/>

The ESD Association <https://www.esda.org/>

Exploring the Characteristics of Field Coupled Crosstalk, Linna and
Semanson, March 31, 2016, In Compliance Magazine, Retrieved from:
<https://incompliancemag.com/article/exploring-the-characteristics-of-field-coupled-crosstalk/>

Fully Differential Amplifiers, James Kahl, Texas Instruments Application
Report SLOA054E, January 2002, revised September 2016 Retrieved from:
<https://www.ti.com/lit/sloa054>

About the Author

Jerry Twomey has designed many consumer, medical, aerospace, and commercial products. This includes: high speed data communication, satellite chipsets, medical instrumentation, cell phones, RF devices, and others. His focus is non-digital electronics, namely those areas where the biggest challenges are. With a breadth of experience designing ICs, PCBs and systems, he has developed the design techniques necessary for reliable electronics.

Jerry has authored multiple trade magazine publications that provide solutions for common electronic design problems. That series of articles became the catalyst for a book that provides a clear-cut methodology to design problem free systems.

Beyond design, Jerry teaches seminars in embedded systems, IC design, EMC, and medical technology. Positions held include: Chair IEEE SD Solid State Circuits, Lecturer UCSD ECE Graduate Studies, Chair IEEE SD Microwave Theory and Techniques, Reviewer IEEE Journal of Solid State Circuits, Reviewer IEEE Journal of Microwave Theory and Techniques, and Instructor UCSD Extension.